False Positives and Shallow Eclipsing Binaries in Transiting Exoplanet Surveys

Thesis

False Positives and Shallow Eclipsing Binaries in Transiting Exoplanet Surveys

Pamela Mary Rowden BA BSc

Supervisor: **Dr Ulrich Kolb,**
Prof Andrew Norton

School of Physical Sciences
The Open University

This dissertation is submitted for the degree of
Doctor of Philosophy

February 2019

Abstract

PLATO (PLAnetary Transits and Oscillations of stars), ESA's M3 mission, is due to launch in 2026. It will aim to detect transits by exoplanets around bright nearby stars and, where possible, to characterise those stars using asteroseismology.

With 24 cameras arranged in four overlapping groups, PLATO will have an excellent signal to noise ratio, making the detection of an Earth-like planet in an Earth-like orbit around a Sun-like star a real possibility. However, the pixel size is large: at 15", there is an increased risk of blends with background eclipsing binaries. This work aims to quantify that risk.

Data from *Kepler* has been used to calibrate the distribution of planets and eclipsing binaries detectable in a transiting exoplanet survey using the population synthesis code, `BiSEPS`. The calibrated synthetic populations have then been used to predict the numbers of exoplanets and eclipsing binaries mimicking planets that PLATO is likely to detect. Other forms of false positive, such as instrumental effects and stellar variability, are beyond the current scope of this project.

My work offers insights into short period ($P < 10$ day) eclipsing binaries detectable in transiting exoplanet surveys, both in terms of initial mass ratio distribution and initial period distribution.

From confirmed *Kepler* planets in *Kepler* Data Release 25, I derive two intrinsic exoplanet distributions which bracket the likely true distribution. These distributions converge at planet radius $0 < log\ R/\mathrm{R}_\oplus < 0.2$.

For the two proposed Long Look fields from Rauer et al. (2014), I find that more exoplanets are likely to be detected in the Southern field, while more blended eclipsing binaries are likely to contaminate observations of the Northern field.

My methods can be extended to other transiting exoplanet surveys, by incorporating the detection parameters of the relevant observatory into the code.

In memory of my mother
Sabina Blanche Lilias Doust
10 November 1925 to 2 January 2011
and her sister, my aunt,
who died during the course of my studies
Margaret Elizabeth Macleod
2 December 1927 to 21 June 2016

Acknowledgements

I would like to gratefully acknowledge the support of my supervisors, Dr Ulrich Kolb and Prof Andrew Norton; the helpful advice from Dr Robert Farmer and Enda Farrell; and the support of my family, without which this thesis would not have been possible: my husband, Nicholas Rowden, our children, Dr James Rowden, Dr Stephen Rowden and Christina Rowden, our daughter-in-law Dr Li Nana, and last but not least my father, Roy Doust, who is always encouraging me to achieve my potential.

Table of contents

List of figures

List of tables

Nomenclature

Other Symbols

B Blue passband, centred on 438 nm, full width half maximum of 97 nm

CCH Ethynyl radical

CH_3OH Methanol

CN Cyanide

CO_2 Carbon dioxide

CO Carbon monoxide

CS Carbon monosulphide

*D*51 A specialist passband, centred on 514 nm, full width half maximum of 16 nm

g' Sloan green passband, centred on 468 nm, full width half measure 137 nm

g mode Gravity wave within a star

H H passband, centred on 1630 nm, full width half maximum of 307 nm

HCN Hydrogen cyanide

HCO^+ An aldehyde (ionised)

HNC Hydrogen isocyanide

I Infrared passband, centred on 857 nm, full width half maximum of 232 nm

i' Sloan near infrared passband, centred on 766 nm, full width half measure 149 nm

J J passband, centred on 1210 nm, full width half maximum of 319 nm

K K passband, centred on 2190 nm, full width half maximum of 390 nm

Kp *Kepler* passband, 420-900 nm

L L passband, centred on 3450 nm, full width half maximum of 472 nm

l Latitudinal co-ordinate in stellar oscillations

n Radial co-ordinate in stellar oscillations

N_2H^+ Diazenylium (ionised)

p mode Acoustic wave within a star

R Red passband, centred on 670 nm, full width half maximum of 207 nm

r' Sloan red passband, centred on 646 nm, full width half measure 99 nm

SO Sulphur monoxide

U Ultraviolet passband, centred on 357 nm, full width half maximum of 66 nm

u' Sloan ultraviolet passband, centred on 365 nm, full width half measure 31 nm

V Visual passband, centred on 547 nm, full width half maximum of 89 nm

z' Sloan infrared passband, centred on 956 nm, full width half measure 230 nm

Acronyms / Abbreviations

2MASS Two Micron All Sky Survey

ABC Approximate Bayesian Computation

ADC Analogue to digital conversion

ADU Analogue to digital unit

AM CVn AM Canum Venaticorum star: a rare type of cataclysmic variable

AMR Age Metallicity Relation

`BiSEPS` Binary Stellar Evolution Population Synthesis

BoL Beginning of life

CCD Charge coupled device

CFHT Canada France Hawaii Telescope

CHEOPS CHaracterising ExOPlanets Satellite

CNO Carbon-Nitrogen-Oxygen cycle

CoRoT COnvection ROtation and planetary Transits

CV Cataclysmic variable

DLR German Aerospace Centre (Deutsches Zentrum für Luft- und Raumfahrt)

DSS Digitised Sky Survey

EMC Electromagnetic compatibility

EoL End of life

ESA European Space Agency

ExoFOP Exoplanet Follow-up Observing Program

`FEE` Front End Electronic

`GUMS` Gaia Universe Model Snapshot

HAT Hungarian Automated Telescope

HATSouth Hungarian Automated Telescope South

IDEM Inverse Detection Efficiency Method

IMF Initial Mass Function

IMR Initial Mass Ratio

IRAM Institut de Radioastronomie Millimétrique

KEBC *Kepler* Eclipsing Binary Catalogue

KELT Kilodegree Extremely Little Telescope

KIC *Kepler* Input Catalogue

KOI *Kepler* Object of Interest

L1 First Lagrangian point

L2 Second Lagrangian point

LLN Long Look North

LLS Long Look South

M1 ESA's first medium class mission

M2 ESA's second medium class mission

M3 ESA's third medium class mission

MAST Barbara A. Mitulski Archive for Space Telescopes

`MESA` Modules for Experiments in Stellar Astrophysics

MES Multiple Event Statistic

`MOST` Microvariablity and Oscillations of STars

NASA National Aeronautics and Space Administration

NExScI NASA Exoplanet Science Institute

NGTS Next Generation Transit Survey

P1 PLATO priority population 1

P2 PLATO priority population 2

P3 PLATO priority population 3

P4 PLATO priority population 4

P5 PLATO priority population 5

PDF Probability Density Function

PLATO PLAnetary Transits and Oscillations of Stars

p-p Proton-proton chain

PSF Point spread function

RV Radial velocity

SED Spectral Energy Distribution

SFR Star Formation Rate

SNR Signal to noise ratio

SuperWASP Super Wide Angle Search for Planets

SWEEPS Sagittarius Window Eclipsing Extrasolar Planet Search

TAMS Terminal Age Main Sequence

TCE Threshold Crossing Event

TESS Transiting Exoplanet Survey Satellite

TRAPPIST TRAnsiting Planets and PlanetesImals Small Telescope

TrES Trans-Atlantic Exoplanet Survey

`TRILEGAL` TRIdimensional modeL of thE GALaxy

TTV Transit timing variation

UKIDSS UKIRT Infrared Deep Sky Survey

UKIRT United Kingdom InfraRed Telescope

WASP Wide Angle Search for Planets

WFIRST Wide Field Infrared Survey Telescope

WISE Wide-field Infrared Survey Explorer

WMAP Wilkinson Microwave Anisotrophy Probe

ZAMS Zero Age Main Sequence

Chapter 1

Introduction

"Are we alone in the Universe?" is a question that humanity has long sought the answer to. The transiting exoplanet survey PLATO 2.0, due to launch in 2026, has probably the best chance of finding an "Earth twin": an Earth-like planet in an Earth-like orbit around a Sun-like star. My work as part of the PLATO 2.0 consortium is to assist in constraining the false positive rate due to extrinsic astrophysical sources. This thesis will describe how I set about this task, and will present my conclusions.

1.1 Planets: from our solar system to the wider Galaxy

Throughout recorded history, people have been interested in the objects that filled the night sky: the Moon, stars and planets. Five planets are observable by the naked eye: Mercury, Venus, Mars, Jupiter and Saturn. In the clear skies of a pre-industrial society, these five specks of light would have stood out: the stars, Moon and Sun always proceed in the same direction across the sky, but these five planets would sometimes move in the opposite direction. This could be explained in a heliocentric system, and over the centuries theories supporting this concept were developed, but there was one key problem: the orbit of the Moon could not be explained by every observable body orbiting the Sun, because it orbits the Earth. With the invention of the telescope, and the discovery by Galileo Galilei of the four Galilean moons of Jupiter in 1610, it was gradually accepted that the Solar System was complex in form, and our understanding of planetary orbits moved forward.

As our telescopes improved, so did our knowledge of the composition of the Solar System. Uranus was discovered in 1781, and Neptune in 1846. Predictions of a planet beyond Neptune led to the discovery of Pluto in 1930. Initially thought to be one body, this is now known to be one planet with five moons. In 1992 the first Kuiper Belt object was announced (Jewitt et al., 1992). Further discoveries followed, leading to Pluto being redesignated as a dwarf planet

in 2006, along with Ceres (the largest body in the asteroid belt) and Eris (a trans-Neptunian object): these were joined by Haumea (a trans-Neptunian object) and Makemake (a Kuiper belt object) in 2008.

In the late 20th century, a series of probes visited the other Solar System planets. Perhaps the most famous are the Voyagers with their grand tour of the gaseous planets, leading to the discovery that all the outer planets had ring systems, not just Saturn; that these planets have countless moons; and that these moons are interesting bodies in their own right. We also learned that, despite the ubiquitous nature of life on Earth, conditions on Mars and Venus were such that extant life seems unlikely.

We thought we understood how and where planets formed from a protoplanetary disc. Conditions would be warmest near the star, and volatiles would not condense out until the "snow line" was reached: thus, the giant planets formed in the volatile rich outer Solar System and rocky planets formed further in, and that was where they stayed. Planet migration was discussed, but on a limited basis (Hayashi et al., 1977; Lin and Papaloizou, 1986): most planetary systems, so the assumption went, would be like our own.

Then the first hot Jupiter was found (Mayor and Queloz, 1995). Suddenly the interest in migration accelerated. Lin et al. (1996), Weidenschilling and Marzari (1996), Ward (1997) and Murray et al. (1998) are just some of the authors who engaged with the subject as the exoplanet era began.

Today, thousands of planets are known. Different catalogues give different numbers. Exoplanet.eu (http://exoplanet.eu/catalog/) lists 3,755 (19 March 2018), the NASA Exoplanet Archive (NExSci) (http://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=planets) lists 3,706 (19 March 2018), while the Open Exoplanet Catalogue (http://www.openexoplanetcatalogue.com/) lists 3,504 (19 March 2018).

Various detection methods have been used, each with its own biases. Their relative success is summarised in Table 1.1, with data drawn from Exoplanet.eu (19 March 2018). More detail on the various methods is given in Table 1.2. Some of the dates of earliest detection by a given method predate the discovery of 51 Pegasi b in 1995 (Mayor and Queloz, 1995).

1.2 Transiting exoplanet surveys: past, present and future

The different exoplanet detection methods described in Table 1.1 and Table 1.2 reveal different pieces of information about the planets that are being observed. Some provide information on the minimum mass, $M_P \sin i$, where M_P is the mass of the planet and i is the angle of inclination, in particular the radial velocity method. Only a transit provides information on the planetary radius. A transit can also be used to constrain the angle of inclination of the

Table 1.1 Exoplanet discovery methods, as listed on Exoplanet.eu, 19 March 2018.

Method	Number	Earliest Detection
Pulsar	29	1988
Radial velocity	746	1989
Imaging	93	1995
Primary Transit	2796	1999
Microlensing	71	2004
Astrometry	1	2010
Transit Timing Variations (TTVs)	7	2012
Other	14	2014

Table 1.2 Summary of exoplanet discovery methods: techniques, biases and examples.

Method	Technique	Bias	Example
Pulsar	Small variations in pulsar timings	Small planets at some distance from the pulsar	PSR 1257 12 b, c & d (Wolszczan and Frail, 1992)
Radial velocity	Periodic Doppler shifts in the spectrum of the host star	High mass planets in short period orbits around low mass stars	51 Pegasi b (Mayor and Queloz, 1995)
Imaging	Direct imaging of planets, for example in protoplanetary discs	Young planets at some distance from the host star	beta Pictoris b (Lagrange et al., 2010)
Primary Transit	Stellar flux dips as the planet passes between the observer and the star	Large, short period planets orbiting small radius stars	HD 209458 b (Henry et al., 2000)
Microlensing	Light from a background object is bent around a foreground planet	Small planets may be observed, but the process is random and unrepeatable	OGLE-2005-390L b (Beaulieu et al., 2006)
Astrometry	Observing small changes in position as the star and planet orbit around a common barycentre	High mass planets in wide orbits around low mass stars	HD 176051 b (Muterspaugh et al., 2010)
TTVs	Systematic variations in the times of ingress and egress of a transiting exoplanet	Pairs of planets in orbital resonance	WASP-47 d (Becker et al., 2015)
Other	Various: See Appendix A		

orbit, which constrains more closely the planet's mass, where a mass estimate is available. Only with information on both mass and radius can the density of the planet be determined.

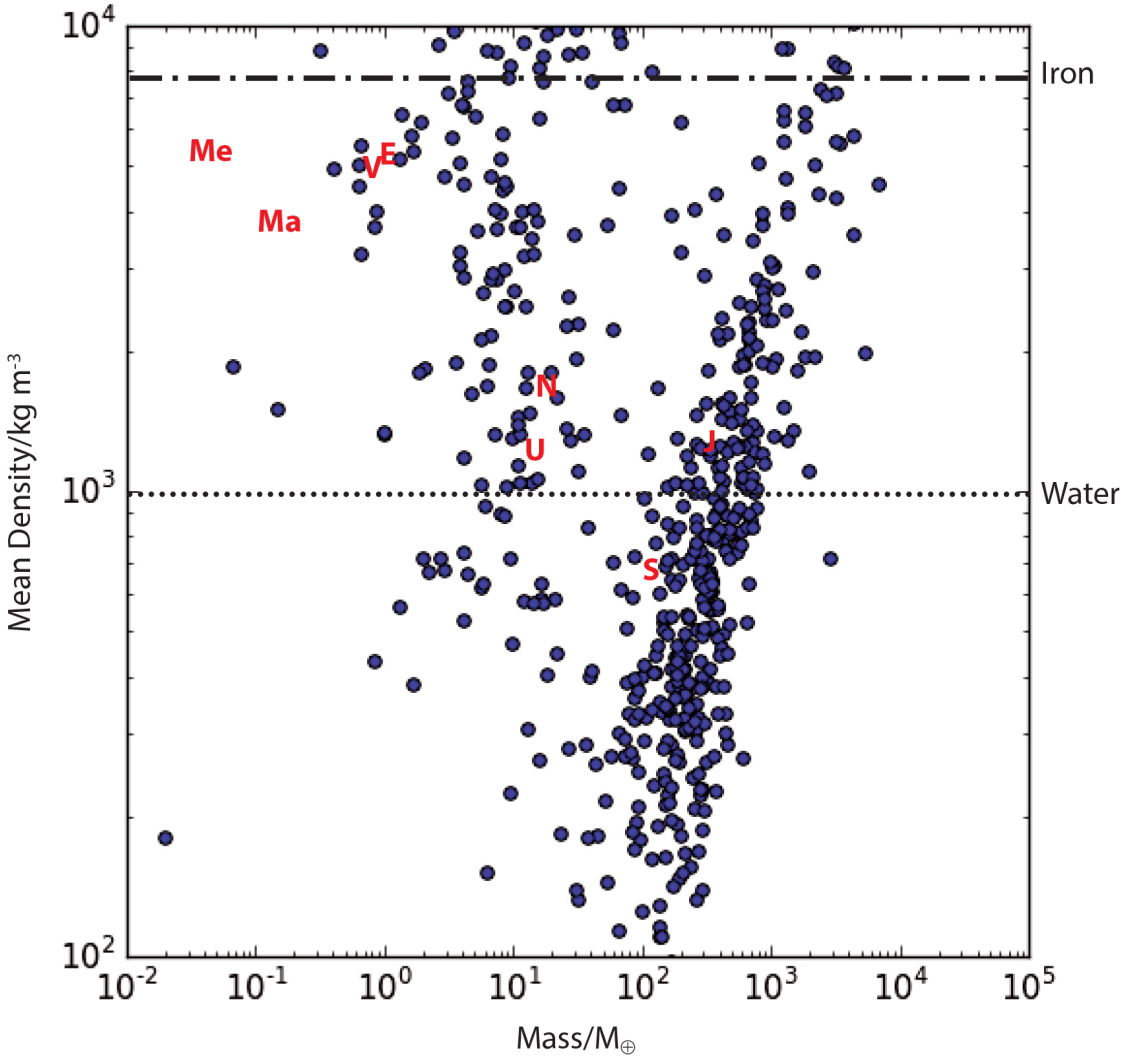


Fig. 1.1 Density v mass, planets listed on exoplanet.eu with both an estimated mass and estimated radius, 5 April 2017. Positions of Solar System planets are indicated by initials. The densities of iron and water are indicated.

Fig. 1.1 plots density against mass for planets listed on exoplanet.eu (5 April 2017) for which information on both mass and radius is available, in the range 10^2-10^4 kg m^{-3}. The densities of water and iron are indicated for reference, and the positions of the planets in the Solar System are indicated by their initials. There are a number of planets with derived

densities greater than solid iron, and this indicates a problem with the estimated mass, the estimated radius, or both: mass and radius are both derived with reference to the host star.

Where the radius of the host star, R_*, and the fractional decrease in flux during a transit, δF, are known, the radius of a planet can be derived using Eq 1.1:

$$R_P = R_* \sqrt{\delta F} \tag{1.1}$$

However, the radius of the star is itself a derived property, often with large error bars. Stellar parameters were updated for *Kepler* DR25 and are both more accurate and more precise than in *Kepler* DR24, and as a result the number of small radius planets has been reduced in favour of intermediate radius planets. This follows high resolution optical spectroscopy of 1305 transiting exoplanet hosts carried out as part of the California-*Kepler* Survey (Petigura et al., 2017). The equation

$$R = \sqrt{\frac{L}{4\pi\sigma T_{eff}^4}} \tag{1.2}$$

relates the stellar radius to stellar temperature (T_{eff}) and stellar luminosity (L), and may in theory be read from a Hertzsprung-Russell (HR) diagram with these axes. However, the error bars tend to be large because T_{eff} and L will often be poorly known, although with Gaia DR2 T_{eff} is expected to be better constrained. T_{eff} may be estimated by observing the lines present in a star's spectrum, a technique used by the California-*Kepler* survey to match observed spectra to T_{eff}, surface gravity log g, metallicity [Fe/H] and radial velocity as a function of inclination $V \sin i$. T_{eff} does not depend on distance d, but L does:

$$L = 4\pi d^2 F \tag{1.3}$$

where F is the out of transit flux. Outside of clusters or associations containing a "standard candle", ie an object which always has the same luminosity, such as a Type Ia supernova, or in the case of a Cepheid variable has a known period-luminosity relationship, distance is almost impossible to estimate without a measurement of stellar parallax, normally only available for the closest stars. All resolvable stars have parallaxes in Gaia DR2, which is of particular benefit in single stars and stars in wide binaries, but from work I am currently undertaking studying unbound companions to *Kepler* Objects of Interest (KOIs), it appears that in the case of faint, distant stars, the parallaxes in Gaia DR2 are poorly constrained. Binaries are expected to be more comprehensively covered in Gaia DR3. Also, an estimate of log g can indicate a star's current evolutionary status, and this, combined with the HR diagram, can provide an estimate of L and hence R.

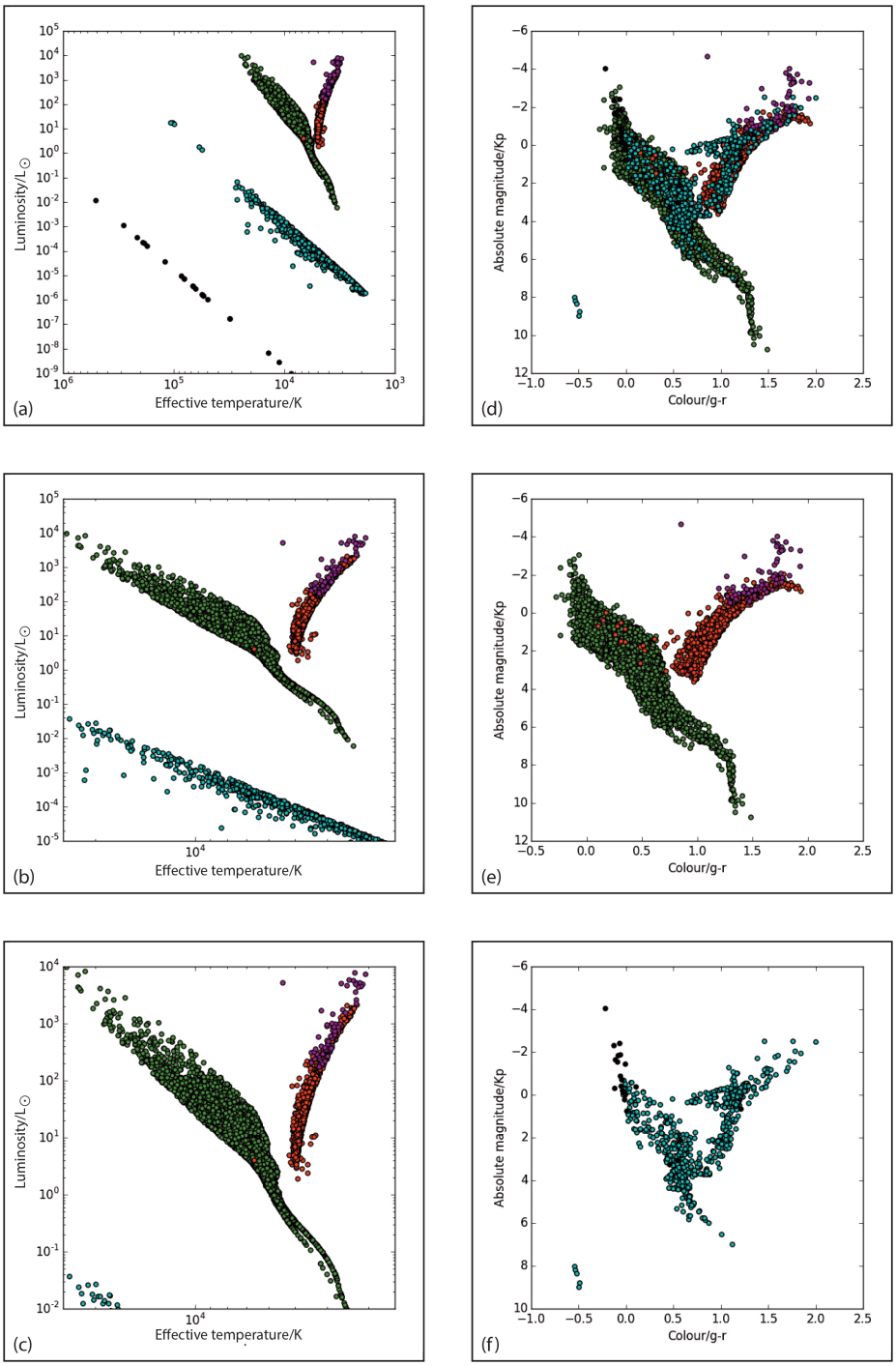


Fig. 1.2 Example Hertzsprung-Russell (HR) diagrams, drawn from `BiSEPS` simulations. Colour code: green, main sequence; red, Hertzsprung Gap and first ascent red giant branch; magenta, core helium burning and asymptotic giant branch; blue, naked helium stars; cyan, white dwarf; black, neutron stars. Panels (a), (b) and (c): $L/\mathrm{L}_{\odot}$ against T_{eff}/K, (a) all stars, (b) centred on $\mathrm{L}_{\odot} \pm 4$, (c) centred on MS stars. Panels (d), (e) and (f): M_{Kp} against colour $(g' - r')$: (d) all systems, (e) excluding compact remnants, (f) compact remnants only.

The HR diagram consists of several well-populated areas, and examples, drawn from a synthetic population generated by `BiSEPS`, are illustrated in Fig. 1.2. Panels (a), (b) and (c) plot L against T_{eff} and include (a) all stars, (b) the region centred on -4 < $log\ L/\mathrm{L}_{\odot}$ < 4 and (c) a focus on the main sequence. Panels (d), (e) and (f) plot absolute magnitude M in the Kp passband against colour $(g' - r')$. The systems are colour coded so that stars on the main sequence are green, Hertzsrpung Gap stars and those on the first ascent giant branch are red, core helium burning and asymptotic giant branch stars are magenta, white dwarfs are cyan and neutron stars are black. Because this plot is based on a synthetic population which covers 1 deg^2, rare stars, such as supergiants, which connect the "top" of the main sequence with the "top" of the first ascent red giant branch, are absent. While the two types of plot both show distinct populations of stars which are not compact remnants, it should be noted that in the colour v magnitude diagrams, compact remnants occupy a similar parameter space to other stars. Therefore, the estimate of surface gravity available from spectroscopy can be essential in distinguishing between compact remnants and other stars in a colour v magnitude diagram. Absolute magnitude M can be found when apparent magnitude m, distance d and extinction A are known via the equation

$$M = m + 5 - 5log_{10}d - A \tag{1.4}$$

Once hydrogen burning has begun, a single star will spend most of its life on the main sequence. The mass of the star will determine the efficiency with which hydrogen burning can occur, and hence the effective temperature and luminosity of the star: more massive stars on the main sequence are hotter and more luminous than low mass stars. The stellar radius is related to the mass, with the exact relation for any given star determined by the principal means of heat transport, both in the core and the envelope. Stars with a mass $\lessapprox 1.1\ \mathrm{M}_{\odot}$ burn hydrogen via the proton-proton (p-p) chain, and have a radiative core, while more massive stars burn hydrogen via the Carbon-Nitrogen-Oxygen (CNO) cycle, and have a convective core. In the envelope, stars with a mass $\lessapprox 0.7\ \mathrm{M}_{\odot}$ will have deeply or fully convective envelopes. More massive stars with p-p chain core processes will have surface convection with a radiative zone between the core and the surface, while stars with core hydrogen burning via the CNO cycle have a fully radiative envelope.

Once the star's radius is known, stellar evolution models may provide an estimate of the star's mass, although spectroscopic measurements where a star is a spectroscopic binary are much more reliable. This is because the mass ratio q, M_2/M_1, is equal to the ratio of velocities, V_1/V_2, around the common centre of mass. Driven by the quest for exoplanets, increasingly small radial velocities are being obtained, with the limit eventually set by the natural variation in the stellar signal.

Knowing the period P and assuming the mass of a planet is negligible compared to the mass of a star, it is possible to estimate a planet candidate's semi-major axis, a, using Kepler's third law. Using the radial velocity amplitude, A_{RV}, the period and eccentricity, e, the planet's mass as a function of the angle of inclination can be found:

$$M_P \sin i = \frac{A_{RV}(M_* + M_P)P\sqrt{(1-e^2)}}{2\pi a} \tag{1.5}$$

Although M_P appears on both sides of Eq 1.5, because $M_P \ll M_*$ the term M_P on the right hand side can be ignored.

So for a system with both RV measurements and a transit, a density can be inferred.

Fig. 1.1 illustrates how useful this can be. Jupiter and Saturn, commonly classed as gas giants, are members of a distinct high mass class with hydrogen/helium (H/He) atmospheres. Neptune and Uranus, also with H/He atmospheres, are also members of a distinct class, known as "Neptunes", sometimes called "ice giants", for which the icy core is a larger proportion of the planet than in the gas giants. The position of Jupiter and Saturn, Neptune and Uranus within their classes is related to their distance from the Sun: hot Jupiters have atmospheres expanded by insolation and hence a lower density. The four rocky Solar System planets have a higher density and are believed to have lost their original H/He atmosphere during the Sun's T-Tauri phase. Mercury's secondary atmosphere is tenuous, Mars and Venus have atmospheres dominated by CO_2 and the Earth's original secondary CO_2 atmosphere has been altered by life and tectonic processes to its present composition.

From Fig. 1.1 the change between gaseous planets with a H/He atmosphere and rocky planets with a secondary CO_2 atmosphere is not immediately clear, and the picture is made no clearer when plotting by density against radius, as in Fig. 1.3. Here, the distribution of gas giants is a clear function of radius, but there is the same "link" between rocky planets and Neptune-class planets. Various authors have studied the nature of planets forming this link, including Adams et al. (2008), Charbonneau et al. (2009) and Southworth et al. (2017), with the very promising detection of H_2O and/or CH_4 in the atmosphere of GJ 1132 b by Southworth et al. (2017): but more data are required. Southworth et al. (2017) speculated from their results that GJ 1132 b may be a "water world". On Earth, H_2O is used by life both as a reactant and a medium in which reactions can take place, while CH_4 may be produced by life, although there are other, geochemical sources.

Following its revision of stellar parameters in 1305 *Kepler* transiting planet hosts, the California-*Kepler* Survey has identified an "evaporation gap" $1.5 < R < 2.0$ $R_\oplus$ (Fulton et al., 2017). This "evaporation gap" is believed to represent a boundary between terrestrial planets and those with a sizeable H/He atmosphere. This gap is not immediately evident in Fig. 1.3, and it should be noted that *Kepler* planets are likely to have short periods, so the conclusions

drawn from this sample may not be applicable to the whole exoplanet population. However, the theory that a proportion of rocky planets relatively close to the host star are the exposed cores of planets whose H/He atmosphere has been evaporated by stellar action may still be valid, with its obvious implication of potentially different formation paths for terrestrial planets close to the host star and those in Venus, Earth or Mars-like orbits.

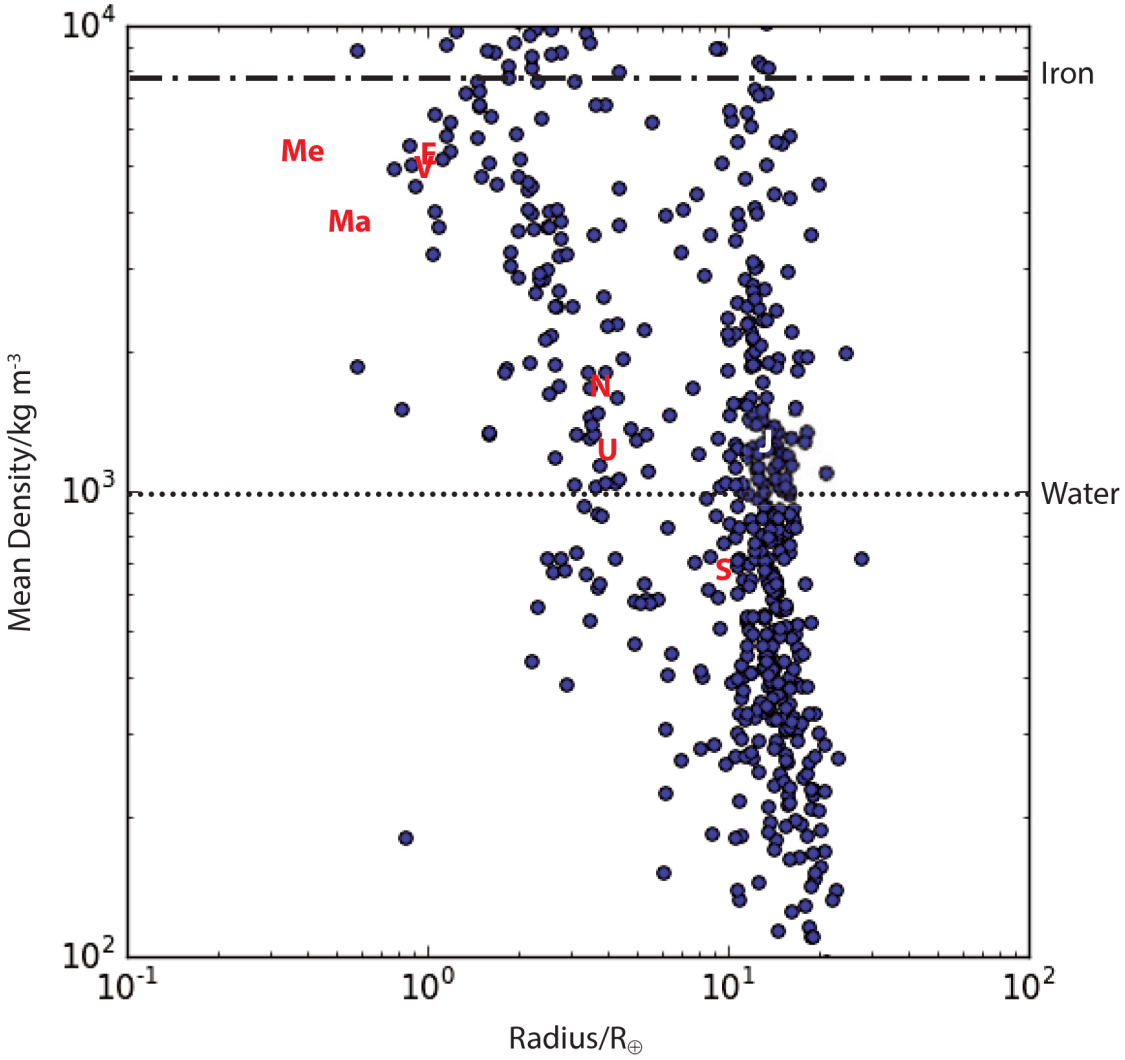


Fig. 1.3 Density v radius, planets listed on exoplanet.eu with both an estimated mass and estimated radius, 5 April 2017. Positions of Solar System planets are indicated by initials. “J” for Jupiter is in white to stand out against the planets in the vicinity. The density of iron and water are indicated.

1.2.1 Previous surveys

Because the resources required for spectroscopic detection of RV are more demanding than those for photometric detection of transits, often a transit will be found first, then followed up with spectroscopy: although if a planet is discovered through RV, for example Proxima b (Anglada-Escudé et al., 2016), the search for a transit will often begin. Various surveys, both ground based and space based, have taken place, observing discrete regions of the sky for long periods of time (Liu et al. (2018), Blank et al. (2018), Patapis et al. (2018)).

Ground based surveys are cheaper to run and easier to maintain. When the original instruments reach the end of their operational life, they can be repaired and upgraded for further service. Space based surveys are above the Earth's atmosphere, removing contamination from telluric lines and 'twinkling' of stars, allowing 24 hour monitoring and enabling the detection of smaller planets than is possible in a ground based survey alone, but are very expensive.

A comparison of the results of selected ground based and space based surveys is given in Table 1.3 and Table 1.4.

The first rocky planet to be confirmed, CoRoT 7-b, was discovered by a space based survey (Queloz et al., 2009).

The most prolific transiting exoplanet survey to date, *Kepler* (Borucki et al., 2010), was launched in 2009 by NASA. 2,295 confirmed planets are listed on the NExSci website (accessed 23 November 2017) (Coughlin et al., 2016). These vary in radius from 0.27-77.76 $R_\oplus$, although only six are > 23 $R_\oplus$. Before the second reaction wheel failed, *Kepler* obtained over four years' data on one 105 deg^2 field centred on l = 76.32°, b = 13.5°. Data from the *Kepler* main mission were used in this thesis to derive information about the intrinsic exoplanet distribution and on the intrinsic properties of eclipsing binaries detectable in transiting exoplanet surveys.

When the *Kepler* spacecraft was no longer able to continue its main mission, the K2 mission was designed. Since 2014 a series of fields have been observed for 80 days at a time, resulting in the discovery of 320 confirmed planets (NExSci website, accessed 9 July 2018), ranging in radius from 0.615-13.966 $R_\oplus$, with a further confirmed planet listed with a radius of 34 $\pm$13 R$\oplus$. Nine confirmed planets do not have estimated radii.

Other space based telescopes, launched for purposes other than transiting exoplanet surveys, have also made transiting exoplanet discoveries. After Spitzer's liquid helium supply was exhausted in 2009, it was used, among other projects, to search for exoplanets. Discoveries include 55 Cnc e (Demory et al., 2011), and the telescope is credited with assisting in confirming the seven-planet system TRAPPIST-1, along with the TRAPPIST ground based telescope (Gillon et al., 2017). Spitzer is owned and operated by NASA (Carey et al., 2010).

Table 1.3 Selected Ground Based Transiting Exoplanet Surveys. Source of information: exoplanet.eu and referenced articles. Date obtained: 7 April 2017. Additional planets have been discovered since.

Name & References	Location & no. of telescopes	Discoveries Period/days	Discoveries Radius/$R_\oplus$	Notes
HAT Bakos et al. (2004)	Mount Hopkins, Arizona (5) Mauna Kea, Hawai'i (2)	1.2-10.9	4.73-23.37	62 planets discovered 3 also WASP planets 2 planets with no radii 77.7 & 219.9 days
HATSouth Bakos et al. (2013)	Chile (1) Namibia (1) Australia (1)	0.84-16.25	6.31-20.8	35 planets discovered
KELT Pepper et al. (2007)	Winter Observatory, Arizona	0.97-7.85	12.51-20.8	13 planets discovered
Qatar Exoplanet Survey Alsubai et al. (2013)	New Mexico	1.34-2.88	12.29-17.4	5 planets discovered
TRAPPIST Jehin et al. (2011) Gillon et al. (2017)	La Silla, Chile (1)	1.5-20	0.76-1.13	7 planets discovered ground based observations combined with space based observations (Spitzer)
TrES O'Donovan et al. (2006)	Tenerife (1) Arizona (1) California (1)	1.3-3.6	12.32-19.12	5 planets discovered
WASP & SuperWASP Pollacco et al. (2006)	La Palma (1) South Africa (1)	0.8-11.55	1.87-22.32	131 planets discovered 3 also HAT planets 2 planets with no radii 1297 & 3725 days Only 2 planets < 0.6 R_J
The XO Project McCullough et al. (2005)	Haleakala Maui, Hawai'i	2.6-4.2	10.9-23.2	6 planets discovered

Table 1.4 Selected Space Based Transiting Exoplanet Surveys. Source of information: exoplanet.eu and referenced articles. Date obtained: 7 April 2017. Additional planets have been discovered since.

Name & References	Agency & launch date	Discoveries Period/days	Discoveries Radius/$R_\oplus$	Notes
CoRoT Bordé et al. (2003)	ESA 2006	0.85-95.27	1.67-16.7	31 planets discovered 2007-2015
Kepler (Borucki et al., 2010)	NASA 2009	0.45-1071	0.27-77.76	2290 planets discovered 2009-2013
K2 Howell et al. (2014)	NASA 2009 (see text)	0.68-44.56	0.615-34	131 planets discovered 2014-present

The Hubble Space Telescope was also involved in a seven-day search for exoplanet candidates in a campaign known as SWEEPS (Clarkson et al., 2008), or The Sagittarius Window Eclipsing Extrasolar Planet Search. 16 planet candidates were found, two of which were confirmed through RV measurements. Hubble is also owned and operated by NASA.

1.2.2 Planned Surveys

Several transiting exoplanet surveys are in an advanced state of planning, each of which will build on previous surveys. The only ground based survey in this group, the Next Generation Transit Survey (NGTS) (McCormac et al., 2017) will use telescopes situated in the high Andes to search for intermediate sized planets. The CHaracterising ExOPlanets Satellite (CHEOPS) (Broeg et al., 2013), a space based mission, will search for transits in known RV planet systems. Two other space based missions, the Transiting Exoplanet Survey Satellite (TESS) (Sullivan et al., 2015) and the PLAnetary Transits and Oscillations of stars mission (PLATO) (Rauer et al., 2014) will both focus on bright stars suitable for RV follow-up and asteroseismology.

NGTS will utilise telescopes based at Paranal, Chile, with the aim of observing Neptune-like and superEarth planets. First light was announced in a press release dated 14 January 2015. NGTS is made up of an array of telescopes observing stars from class K to early M with $V \leq 13$ at 600-900 nm. The consortium behind NGTS includes institutions based in the United Kingdom, Switzerland, Germany and Chile. Advantages of Paranal as a location include particularly low water vapour in the atmosphere (Wheatley et al., 2018).

TESS launched in April 2018 and is a NASA mission. It will seek to identify planets through transit photometry orbiting nearby stars. The TESS website states that 200,000 stars will be monitored in a two year all-sky survey, part of a three year science mission, and that the satellite is expected to catalogue > 1,500 planets, of which about 500 would be < 2 $R_\oplus$. Following the all-sky survey, searching 26 overlapping 24° x 96° fields for 27 days/field, the continuous viewing zone at the ecliptic poles will be observed for 350 days (Ricker et al., 2015).

CHEOPS is expected to launch late in 2018, and is an ESA mission. During the three and a half year mission, 500 targets already known to be exoplanet hosts through RV observations are expected to be observed, with the aim of detecting shallow transits through the use of transit photometry. The ESA website describes the mission's main aims as measuring the bulk density of super-Earths and Neptunes orbiting bright stars, and to assist in the selection of suitable targets for further follow-up (Broeg et al., 2013).

PLATO, due to launch in 2026, is an ESA mission which is expected to revolutionise our knowledge and understanding of exoplanet science (Rauer et al., 2014). This thesis is

concerned with part of the preparatory work for this mission, and hence this mission will be considered separately in Section 1.3.

1.3 PLATO 2.0: background and aims

PLATO 2.0 is ESA's M3 mission (Rauer et al., 2014). M3 is the third of ESA's medium class missions, part of ESA's long-term Cosmic Vision programme. PLATO 2.0 falls into theme 1 of the Cosmic Vision, "What are the conditions for planet formation and the emergence of life?"[1]. The "2.0" refers to the fact that PLATO had previously been unsuccessful in competing for the M1 and M2 slots (which were jointly considered), and was revised prior to successful submission for the M3 slot.

PLATO is expected to operate in the L2 position. The total field of view is expected to be 2,232 deg^2. While the final observing strategy has yet to be confirmed, this thesis is using as its basis the strategy proposed in Rauer et al. (2014) of two "Long Look" fields, one centred on $l = 65°$, b = 30°, known as Long Look North (LLN) and the other centred on l = 253°, $b = -30°$, known as Long Look South (LLS). Either each will be observed for two years, or LLS will be observed for three years and LLN for one year. The long look fields are expected to be supplemented at the end of the mission by the "Step and Stare" fields, each of which would be observed for six months, provided the instrument is still working as intended and there is sufficient funding. As the mission's name suggests, the satellite will be recording information both on transits and on stellar oscillations.

L2 is the second Lagrangian point (Fig. 1.4), and PLATO is expected to operate in the L2 position in the Earth-Sun system. The five Lagrangian points are the points where the combined force of gravity from two bodies, such as a star and a planet, and the centrifugal force are balanced. A smaller third body may remain at rest in the co-rotating frame at a Lagrangian point. The Gaia Space Observatory (ESA) operates at L2 in the Earth-Sun system. Spacecraft in an L2 orbit which complete their missions can be moved to other positions, such as the Herschel Space Telescope (ESA) and WMAP (Wilkinson Microwave Anisotropy Probe) (NASA). Satellites expected to operate at the L2 point in the Earth-Sun system in the future include WFIRST (Wide Field InfraRed Survey Telescope) (NASA), expected to be launch in the early 2020s: and PLATO, due to launch in 2026.

1.3.1 PLATO's field of view and cameras

PLATO will have two fast cameras, for studying stars $4 < V < 8$, and four groups of normal cameras, for observing stars $8 < V < 16$. Rauer et al. (2014) stated that there will be eight

[1]http://sci.esa.int/cosmic-vision/46510-cosmic-vision/?fbodylongid=2152

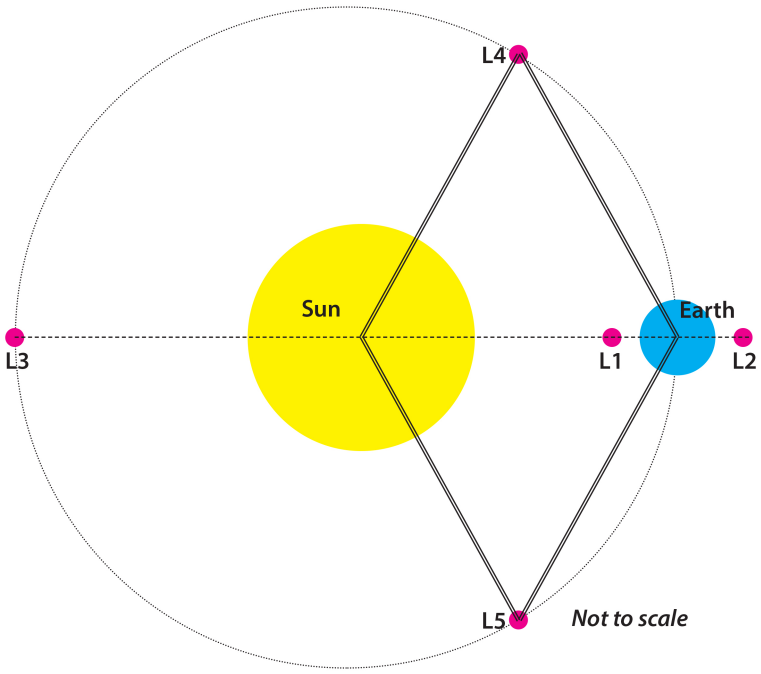


Fig. 1.4 The five Lagrange points in a two-body system. These are the points where the gravitational force and centrifugal force from the two bodies is balanced. PLATO is expected to orbit in the L2 position in the Earth-Sun system. Not to scale.

cameras in each group of normal cameras, but as the project progressed this was reduced to six cameras per group at the official adoption of the mission by ESA in June 2017. The implications of the reduction on PLATO's performance will be examined in Chapters 5 to 8.

Technical details from Table 4.1 on page 71 of the PLATO Definition Study Report, 2017[2] are included here in Table 1.5

Fig. 1.5 shows one configuration illustrating how the PLATO cameras will overlap, assuming the angle between the centre of the global field and the centre of the field of a group of cameras is 0.35, expressed as a fraction of the field, or 9.6°. Instrument design is ongoing, and an alternative configuration may be used in flight.

The large plate scale of 15.0 arcsec/pixel is the reason why the possibility of false positives due to background eclipsing binaries is a particular concern. The PLATO Definition Study Report states that a window around each target star will be formed, known as an "imagette", typically set at 6 x 6 pixels for a normal camera (9 x 9 pixels for a fast camera) to ensure that all information relevant to the point spread function (PSF) of a given star is read and analysed, whether on the ground or on board. This will help verify the quality of the photometric and centroiding data, and will be especially useful in identifying nearby background eclipsing binaries, according to the PLATO Definition Study Report.

[2]http://sci.esa.int/jump.cfm?oid=59252

Table 1.5 Summary of PLATO technical details, drawn from Table 4.1 on p 71 of the PLATO Definition Study Report

Aperture of cameras	120.0 mm
Field of view, fast cameras	≈ 619 deg^2
Field of view, normal cameras	≈ 1037 deg^2
Total field of view	≈ 2232 deg^2
Covered by 4 groups of cameras	≈ 301 deg^2
Covered by 3 groups of cameras	≈ 247 deg^2
Covered by 2 groups of cameras	≈ 735 deg^2
Covered by 1 group of cameras	≈ 949 deg^2
Fast camera CCDS	1 per camera, 4,510 x 2,255 pixels
Normal camera CCDs	4 per camera, 4,510 x 4,510 pixels
Pixels	18 μm square
Plate scale	15.0 arcsec/pixel
CCD temperature	$< -65^\circ$ C, maintained by passive cooling
Spectral range	500-1000 nm
Cycle period for normal cameras	25 s
Exposure time, normal cameras	22 s
Cycle period for fast cameras	2.5 s
Exposure time, fast cameras	2.3 s
Read out frequency	3 Mpx/s
Expected mass of payload	533 kg
Expected power required by payload	≈ 820 W

1.3.2 PLATO priority populations

One of the main aims of PLATO 2.0 is to detect an Earth twin. As shown in Table 1.3 and Table 1.4, the only place to really do this from is space. Of the seven rocky planets orbiting TRAPPIST-1, the first three were detected by TRAPPIST, a ground based survey, and the remaining four through space based observations by Spitzer (Gillon et al., 2017).

Kepler has shown that, when a satellite stares at one area of the sky for a long period of time, it is possible to detect many planets. However, many *Kepler* planet hosts are unsuitable for spectroscopic follow-up as they are too faint, and effective follow-up is essential if an Earth-twin is to be confirmed.

PLATO has set five priority populations, numbered in order of importance. The four that apply to the Long Look phases are given in Table 1.6: P3 is similar to P2, but applies only to the Step and Stare phase. P2 (and therefore P3) is a subset of P1, and P1 is a subset of P5.

The Sun is class G2, so would fall within P1, P2 or P5, depending on the distance of an observer from the Sun and hence the apparent magnitude, m. M class stars, being intrinsically less luminous than FGK stars, have a fainter magnitude limit set.

The earliest stars on the PLATO target list will be stellar class F5. Hotter, more massive stars tend to be more active than Sun-like stars (Pace and Pasquini (2004), Martínez-Arnáiz

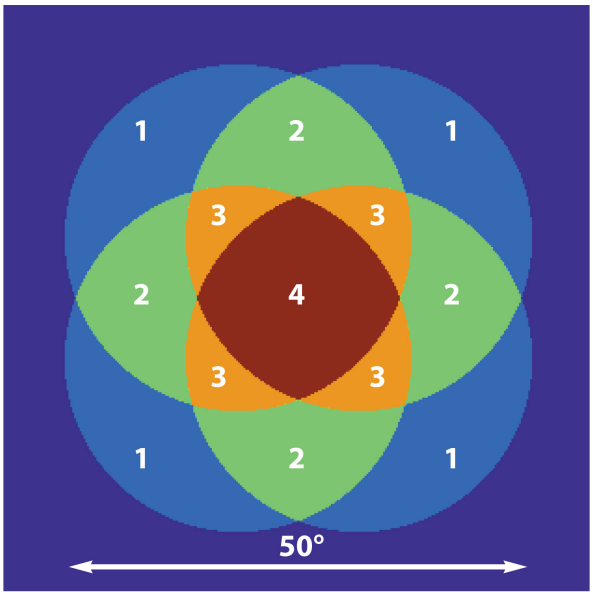


Fig. 1.5 PLATO field of view, illustrating overlapping groups of cameras. This diagram assumes the angle between the centre of the global field and the centre of the field of a group of cameras is 0.35, expressed as a fraction of the field, or 9.6°. Numbers refer to the number of groups observing a given region. This figure was generated using code provided by Claude Catala of l'Observetoire de Paris, and represents one possible configuration. An alternative configuration may be used in flight.

Table 1.6 PLATO populations appropriate to the Long Look fields. P3 is similar to P2, but covers the Step and Stare phase only. Source of data: Rauer et al. (2014).

Priority	**Stellar Class**	**Apparent Magnitude** m
P1	F5 to K7	$V < 11$
P2	F5 to K7	$V < 8$
P4	M	$V < 16$
P5	F5 to K7	$V < 13$

et al. (2010)): stellar activity generates starspots (the extrasolar equivalent of sunspots), which can be confused with exoplanet transits. Early stars are also more massive than late stars, and hence more luminous: the habitable zone of an A-class star or an early F-class star will be further from the star than the Earth is from the Sun. The year of a planet in the habitable zone of such a star will be longer than an Earth year, making it less likely that there will be two detections, let alone three, while a Long Look field is being observed. Early stars, being more massive than late stars, also have a shorter lifespan. For all these reasons, studying stars earlier than class F5 for transiting exoplanets, especially one that is an Earth twin, is not likely to be fruitful.

1.3.3 The Long Look Fields

If the fields proposed in Rauer et al. (2014) are those that are selected for the mission, then the LLS field will be centred approximately on the Pictor constellation and the LLN field will be centred approximately on Lyra and Hercules. The LLN field would contain within its boundaries the field examined during the *Kepler* main mission. Centred on $b = \pm 30^{\circ}$, variation over Galactic latitude in the Long Look fields will be much more significant than variation over Galactic longitude during the observing campaigns. This is due to the geometry of the thin disc and thick disc, described in more detail in chapter 2: in particular, the vertical scale height (from the Galactic plane) is smaller than the radial scale height (from the Galactic centre). In both fields, close to the Galactic plane stars that may be exoplanet hosts will be much more plentiful, but there will also be a greater risk of contamination from eclipsing binaries, both within the pixel in which the target sits and in neighbouring pixels. Closer to the celestial poles, fewer planet hosts will be observed, but there will also be fewer background stars, increasing the probability that a detected signal is not a false positive. One consideration under investigation is whether it would be possible to move the fields up to 5° closer to the Galactic plane, depending on the level of contaminants (Rauer et al., 2014): moving closer than this to the Galactic plane for the Long Look fields is ruled out by the need to look above/below the Sun at appropriate points in the orbit. This restriction would not apply to the same extent to the Step and Stare fields, as these would only be observed for six months/field.

As shown in Fig. 1.5, the best coverage will be at the centre of the field of view, where all four groups of cameras overlap.

1.3.4 Asteroseismology

As discussed in Section 1.2, and shown in Eq 1.1 and Eq 1.5, to obtain an accurate radius and mass for an exoplanet, it is essential to know the mass and radius of the host star. Rauer et al.

(2014) stated that one of the aims of PLATO 2.0 is to use asteroseismology to confirm stellar radii to within $\pm$1-2% and to confirm stellar masses and ages to within $\pm$10%.

Asteroseismology measures the oscillations in a star to obtain information about its internal structure. There are two possible modes, p modes and g modes, where p modes are trapped acoustic waves and g modes are trapped gravity waves. In a spherical co-ordinate system, oscillations within a star depend on the radial co-ordinate and the polar co-ordinate rotating around the axis of rotation, with no dependence on the azimuthal co-ordinate. The radial co-ordinate is given as n and the polar co-ordinate as l in the relevant notation.

In the second appendix of Rauer et al. (2014), it is noted that PLATO is unlikely to observe modes greater than $l = 3$, but that the l co-ordinate is useful in providing information on the internal angular velocity of a star. The stated aim is to achieve with PLATO the same resolution as is possible with the Sun, namely an accuracy of $\approx 0.1\mu$Hz. Where ν is frequency, there are two separations, one large Δ_l

$$\Delta_l = \nu_{n,l} - \nu_{n-1,l} \tag{1.6}$$

between modes of the same degree l and the other small Δ_{02} or Δ_{01},

$$\Delta_{02} = \nu_{n,0} - \nu_{n-1,2} \tag{1.7}$$

$$\Delta_{01} = \nu_{n,0} - \frac{\nu_{n-1,1} + \nu_{n,1}}{2} \tag{1.8}$$

where Δ_{02} is between modes $l = 0, 2$ and Δ_{01} is between modes $l = 0, 1$. The mean stellar density can be obtained from Δ_l, since it measures the acoustic travel time from the stellar core to the stellar surface, which will assist in constraining the mass and radius of the star.

Δ_{01} and Δ_{02} provide information on the internal structure of the star, which is useful in constraining the age of the star. This information will be obtained both for stars which are and are not known to be planet hosts. Of course, the absence of known planets does not mean that a star does not host planets: indeed, it is often speculated that all or most stars host planets, it is just that not all planetary systems are detectable from Earth.

To properly understand the evolution of planetary systems over time, especially in the context of migration, understanding the age of observable systems is considered essential. However, dating single stars is not easy, especially as stars spend a long time on the main sequence. During this time, although the star's radius will expand over time and mass will be lost in stellar winds, these changes are slow and slight compared to more dramatic changes in later life, such as crossing the Hertzsprung-Russell gap, ascending the Red Giant Branch and ejecting planetary nebulae. For a star which is well settled on the main sequence, particularly

a low mass star with a very long main sequence life time, determination of the exact age through a colour versus magnitude diagram, for example, can be problematic.

One method of dating stars, based on the fact that stars appear to spin fastest at birth (Chanamé and Ramírez (2012), García et al. (2014)), slowing down as they evolve through the main sequence, is gyrochronology. However, it has been observed that some exoplanet hosts are spinning faster than would be expected when their ages are estimated through other means, such as the use of isochrones in the HR diagram (Lanza (2010), Brown et al. (2011a)).

Gyrochronology has the advantage of being model independent, based on a period/age/colour relation (Brown (2014) and references therein). Awareness of an issue involving exoplanet hosts arose in Lanza (2010), a study of hot Jupiters and their host stars. This found that stars with $T_{eff} \geq 6000$ K and a rotation period $P_{rot} \leq 10$ days had a P_{rot} that was equal to or double the orbital period of the planet, and that stars with $T_{eff} \leq 6000$ K and/or $P_{rot} \geq 10$ days showed a trend towards synchronisation with increasing T_{eff} and/or decreasing P_{rot}.

Brown (2014) and Maxted et al. (2015) have both examined the question of whether the discrepancy between ages derived from gyrochronology and ages derived from isochrones in exoplanet hosts is a result of tidal interactions between planet and star, a natural consideration arising out of the study by Lanza (2010). Brown (2014) studied 68 planet hosts, finding that isochrones tended to give slightly older ages than gyrochronology but that this was not correlated to the tidal timescale. Maxted et al. (2015) found that in a sample of 28 stars, the gyrochronological age was significantly lower than the isochronal age in about half the stars in the sample and that, while tidal interactions between the star and planet were a reasonable explanation for the discrepancy in some cases, this did not apply to all cases, and that there was no clear correlation between the gyrochronological age and the strength of the tidal force of the star on the innermost planet. Both Brown (2014) and Maxted et al. (2015) find that in some cases isochronal ages may overestimate a star's age. Jackson et al. (2009) had previously discussed the possibility of tidal disruption of planets as one explanation for the observed "three-day pile-up" of hot Jupiters, with the remnants of tidally disrupted planets presumably either falling into the star or surviving as fragments of the original giant planet. Presumably, if this had happened, there would have been previous tidal interactions that could have affected a star, but the planet that was the direct cause of those interactions would no longer exist in its previous form.

The information on the internal structure of a star from asteroseismology may help resolve this issue. Over the main sequence lifetime, hydrogen is converted into helium in the stellar core. The proportion of helium to hydrogen in the core is therefore a guide to the stellar age, in conjunction with the star's effective temperature: hotter, more massive stars convert hydrogen to helium at a higher rate than cooler, less massive stars, and hence have shorter main sequence lifetimes. The angular velocity in the star's interior can be related to the estimated age through

gyrochronology. Combining these sources of information with the information that can be extracted through more traditional means, such as measuring the rotational velocity through spectroscopy, should throw light on why some planet hosts appear younger when their age is assessed through gyrochronology than through the use of isochrones. The enhanced accuracy of PLATO's transit observations may also allow the observation of the remnants of tidally disrupted planets, assuming the angle of inclination is favourable, and if this is the case, another piece will have been added to the puzzle of dating stars accurately.

1.4 Stellar formation: binaries and higher hierarchies

Stars form when a molecular cloud collapses, often into clumps, with each clump potentially becoming an individual star. For this reason, stars will often form in clusters. A star forms when the mass in a clump exceeds the Jeans mass (Jeans, 1902). The Jeans mass

$$M_{Jeans} \equiv \frac{3kT}{2G\overline{m}}R \tag{1.9}$$

is the minimum mass of a molecular cloud for which gravitational collapse is possible. T is the temperature, R the radius, $\overline{m}$ the mean mass of the particles making up the cloud, k is Boltzmann's constant and G is the gravitational constant.

Within a given volume, the Jeans mass may be exceeded either by increasing the temperature or by reducing the mean molecular mass. In most cases, increasing the temperature, for example as the result of a shock wave from a supernova, is more likely than changing the mean chemical composition. Molecular clouds consist mostly of molecular hydrogen (H_2), with small amounts of other molecules. For example, Gratier et al. (2017) observed the Orion B Cloud with the IRAM 30 m single dish telescope. In addition to H_2, the following molecules were identified spectroscopically: ^{12}CO, ^{13}CO, CS, HCN, HCO^+, SO, HNC, CCH, $C^{18}O$, N_2H^+ and CH_3OH. As well as a background dominated by H_2, Gratier et al. (2017) were able to identify regions of UV illumination, with a positive contribution from CCH, CN and anti-correlated with N_2H^+ and CH_3OH, and dense cores with a positive contribution from N_2H^+ and CH_3OH and a negative contribution from ^{12}CO and ^{13}CO.

These results suggest that star formation has already started in the Orion B cloud. The dense cores are likely to be protostars still in the process of contracting, while the regions of UV illumination mark the location of young stars already on the main sequence.

With one cloud giving birth to many stars, binaries and higher hierarchies will form where stars are gravitationally bound together. This occurs when kinetic energy equals gravitational potential energy within a system, where the system consists of two or more stars. The common centre of mass is the barycentre.

In a binary system, two stars labelled A and B orbit the barycentre: and this is the basis of all other hierarchies. As an analogy, imagine the solar system consisted only of three bodies, the Sun, the Earth and Moon. The Earth and Moon orbit one another, while the Earth-Moon system and the Sun also orbit one another. Allowing for the rather obvious difference in mass, the Earth-Moon system may be compared with the inner pair of a triple system, labelled Aa and Ab, while the Earth-Moon and Sun system may be compared with the outer pair of a triple, labelled (Aa+Ab) and B. A quadruple system may consist of two inner pairs, for example Aa and Ab, Ba and Bb, and an outer pair, (Aa+Ab) and (Ba+Bb), or may have an inner pair, Aaa and Aab, an intermediate pair, (Aaa+Aab) and Ab, and an outer pair ((Aaa+Aab)+Ab) and B. Some possible configurations are illustrated in Fig. 1.6. Any system observed as a binary may be part of a higher hierarchy.

KIC 4150611 (KOI 3156) is a good example of a multiple system and of the interesting non-planetary objects that can be observed in a transiting exoplanet survey. Studied by Hełminiak et al. (2017), it was found by *Kepler* to have four possible transits: at 94.2 days, 8.65 days, 1.52 days and 1.43 days respectively. Hełminiak et al. (2017) found that this one object in the *Kepler* Input Catalogue could be resolved spectroscopically to three point sources, which the authors labeled A, B and C. Hełminiak et al. (2017) concluded that C is most likely a background eclipsing binary, with a period of 1.43 days, and that A and B form a quintuple system of a similar configuration to the orange example in Fig. 1.6. The outermost pair, A and B, have an estimated orbital period of 1,000 years. B is a pair of G class stars, Ba and Bb, with a period of 8.65 days. A has an inner pair and an outer pair: the outer pair, Aa and Ab, has a period of 94.2 days. The inner pair, Ab1 and Ab2, is made up of two class K or M stars with a period of 1.52 days. Aa is class F1 and is described as a hybrid pulsator, as it demonstrates δ Scuti type pulsations with the highest peak at 20.243 d^{-1}, and γ Doradus type pulsations with the highest peak at 2.6064 d^{-1}, and therefore is an interesting object in its own right.

This system demonstrates the difference between foreground and background eclipsing binaries as well as the effect of blending, and demonstrates the importance of follow up observations in confirming that a candidate is a planet and not an eclipsing binary.

1.5 Eclipsing binaries

In a binary system, if one star passes between an observer and the other star in the system, an eclipse occurs. Calculating the radius of the eclipsing star uses the same principles as those used to calculate the radius of a transiting exoplanet, but allows for the fact that stars are self-luminous while planets are not.

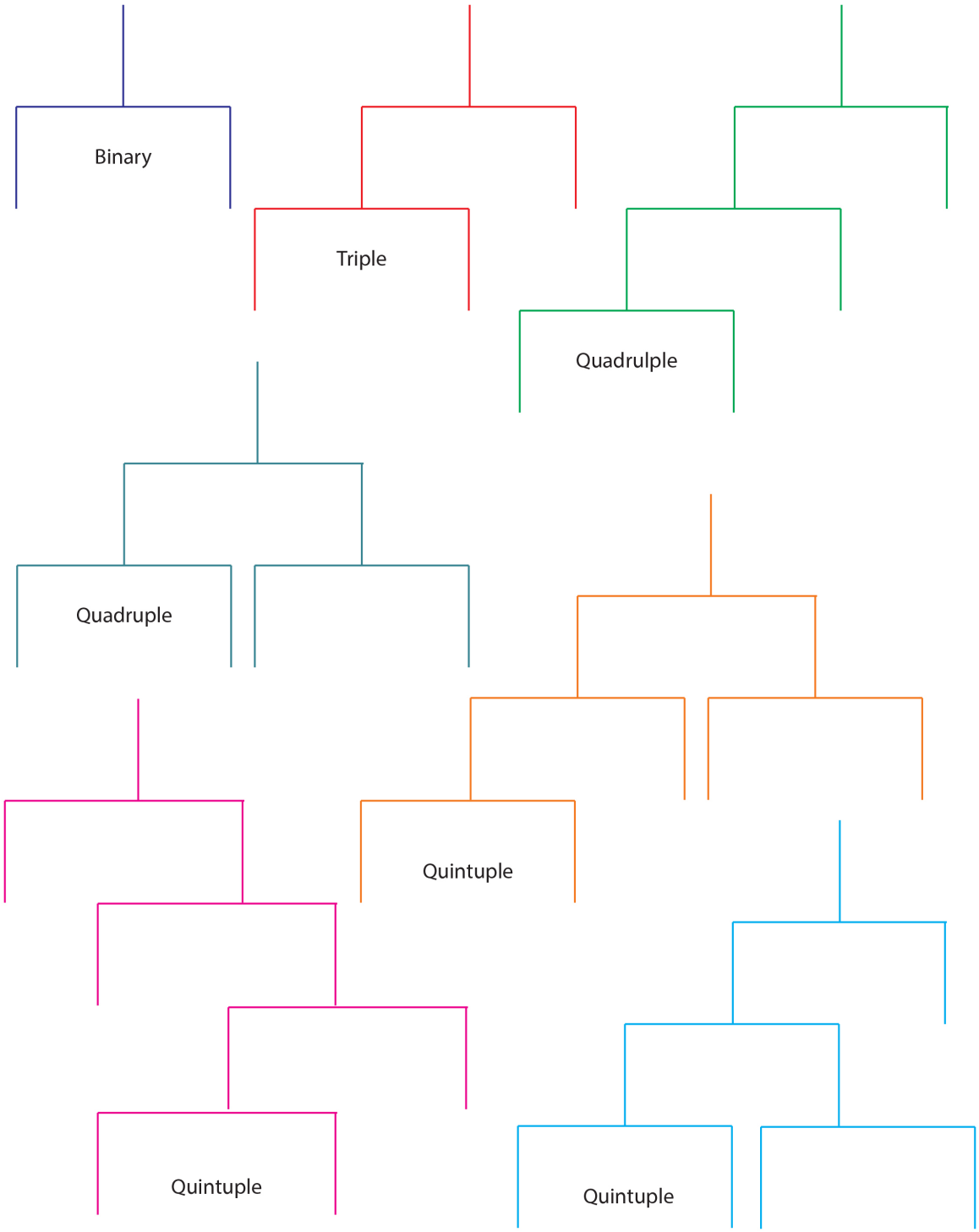


Fig. 1.6 A mobile diagram showing possible configurations of binaries and higher hierarchies.

$$\Delta F = \frac{F_2 + (F_1 - (A_2/A_1)F_1)}{F_1 + F_2} \tag{1.10}$$

where ΔF is the normalised in transit flux, F_1 is the flux of the body being eclipsed (star 1), F_2 is the flux of the eclipsing body (star 2), A_1 is the area of star 1 and A_2 is the area of that part of star 2 which occludes part of star 1, and is only a circle when the whole of star 2 passes in front of star 1. $A_2/A_1 \leq 1$ must be true even in the extreme case of a smaller body totally eclipsed by a larger one, in which case the expression would reduce to

$$\Delta F = \frac{F_2}{F_1 + F_2} \tag{1.11}$$

Since the normalised fractional eclipse depth is

$$\delta F = 1 - \Delta F \tag{1.12}$$

ie the normalised out of transit flux minus the normalised in transit flux. Similar principles apply to determining δF for both planets and eclipsing binaries: the key difference is that in a binary system both the occluding and occluded bodies are self-luminous, but when a planet transits a star only the star is self-luminous.

Eq 1.10 is, of course, a simplification: limb darkening also needs to be taken into account for a full description. Care also needs to be taken to distinguish between the primary and secondary eclipses, where both are observable, as δF will be different in these two cases and the subscripts 2 and 1 will refer to different stars in each case (the smaller star is star 2 in the primary eclipse and star 1 in the secondary eclipse, for example). However, if the ratio of areas is known, the ratio of radii may also be known (depending on the impact parameter and hence the proportion of star 2 which occludes star 1), and if the flux of the two stars can be distinguished and the colour of at least one of the stars is known, an estimate of the radii of the stars can be obtained.

Transiting exoplanets surveys will in the first instance estimate the radius of an eclipsing or transiting body with the assumption that it is not self-luminous. Where the calculated radius is clearly stellar, or the light curve is highly sinusoidal, such systems will not be confused with a transiting exoplanet. However, there are three cases where the eclipse depth may be shallow enough for confusion to occur:

1. A grazing eclipse.
2. Where the eclipsing body is a white dwarf.
3. Where the eclipse depth is diluted through blending.

The *Kepler* main mission has produced a considerable amount of data on eclipsing binaries, with 2,876 systems listed in the *Kepler* Eclipsing Binary Catalogue[3] (18 April 2017). This is greater than the list of confirmed planets, which stands at 2,290 (18 April 2017). The binaries in the *Kepler* Eclipsing Binary Catalogue are all cases where the target is an eclipsing binary. Planets, background eclipsing binaries and cases of stellar activity can be searched for as individual items. However, one of the fields returned, "InCat" (short for "In Catalogue?") will carry the designation of "False", rather than "True", when the item searched for is not an unblended eclipsing binary.

The light curves of binary systems can be very different from one another, depending on their morphology. For that reason, the *Kepler* Eclipsing Binary Catalogue includes a morphology parameter in the range 0-1, best described in Matijevič et al. (2012). As discussed in Section 7.3, as the morphology parameter increases, ellipsoidal variations become more evident and eclipse duration as a fraction of orbital period increases. However, there is no formal link between these two observable characteristics and the morphology parameter (Andrej Prsa, priv. comm.). In Matijevič et al. (2012) examples of light curves with a given morphology parameter are included, as are definitions of the expected nature of binaries with a given morphology parameter, described in more detail below. Matijevič et al. (2012) emphasised that the morphology parameter does not specify a given type of binary, it simply indicates the most likely type, and that is the formalism followed here.

Binary light curves can be classified in four ways: detached, semi detached, contact or overcontact, and ellipsoidal. The morphology parameter as utilised by Matijevič et al. (2012) describes the *probable* classification of individual binary systems, and is not intended as a hard-and-fast description.

Detached: Matijevič et al. (2012) designated detached binaries as having a morphology parameter of < 0.5. As the name suggests, detached binaries are systems which are not in physical contact with one another. These stars tend to have longer orbital periods and are generally spherical in shape. KIC 9027841 and KIC 3113266, on the bottom row of Fig. 1.7, with morphology parameters appropriate for detached systems, might easily be mistaken for a transiting exoplanet were it not for the distinctive 'V' shape of the primary eclipse, combined with the presence of a secondary eclipse.

Semi-detached: As a star ages, in particular as it moves across the Hertzsprung Gap and then up the Red Giant Branch, its radius will increase. When stellar expansion reaches the L1 point (Fig. 1.4), possibly as a result of tidal bulging, the plasma on the surface of the star at L1 will be equally under the gravitational influence of both stars. Expanding beyond L1 means that, along the line joining the two stars, matter from the expanding star will be gravitationally

[3]http://keplerebs.villanova.edu/

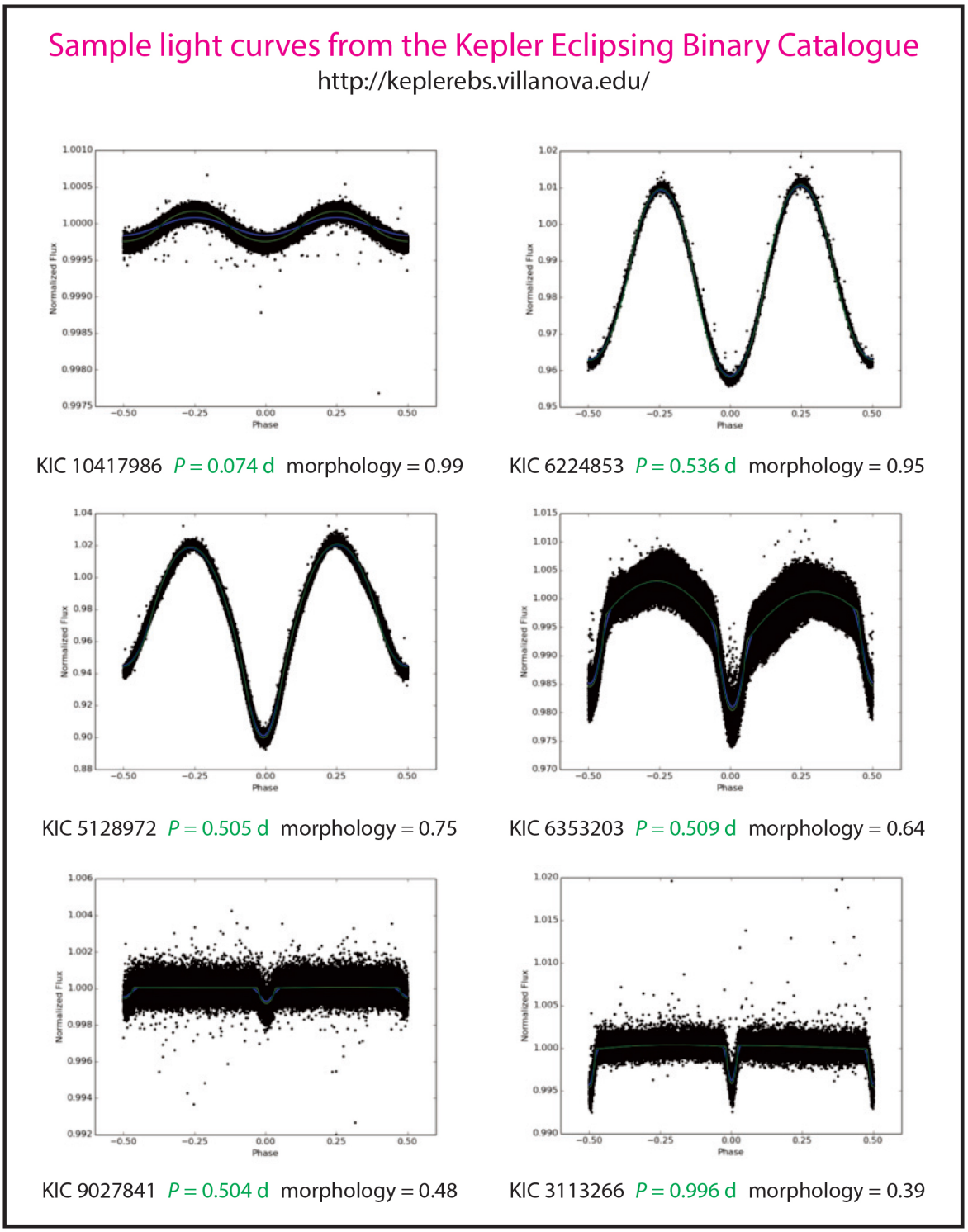


Fig. 1.7 Sample light curves from the *Kepler* Eclipsing Binary Catalogue, $P < 1$ day. Top row: ellipsoidal binaries. Middle row: contact binary (left) and semi-detached binary (right). Bottom row: detached binaries. See text for further details on the different types of binaries.

attracted to the companion, and mass transfer will occur. The first star has filled its Roche lobe, and the system is semi-detached.

Matijevič et al. (2012) designated semi detached binaries as having a morphology parameter of $\approx$ 0.5-0.7, and in Fig. 1.7 KIC 6353203 is such a system. Its period is similar to that of KIC 6224853 (ellipsoidal) and KIC 9027841 (detached), and its light curve carries characteristics of both, with clear eclipses at $\phi = 0$ and $\phi = 0.5$, with the eclipse depth at $\phi = 0.5$ distinctly shallower than at $\phi = 0$, but with a sinusoidal appearance at $\phi = 0.25$ and $\phi = 0.75$.

Contact or overcontact: Where both stars fill their Roche lobes, the surfaces of the two stars are in physical contact. This type of system is referred to both as a contact or overcontact system, and can result in the stars merging in a common envelope phase. Matijevič et al. (2012) designated contact binaries as having a morphology parameter of $\approx$ 0.7-0.8, and in Fig. 1.7 KIC 5128972 is such a system. As with KIC 10417986 and KIC 6224853 (both ellipsoidal), the light curve of KIC 5128972 appears sinusoidal in nature and it is difficult to make out an actual eclipse, but the minima at $\phi = 0$ and $\phi = 0.5$ are distinctly different.

Ellipsoidal: Matijevič et al. (2012) designated ellipsoidal binaries as having a morphology parameter of > 0.8, and two examples are shown in the top row of Fig. 1.7. Ellipsoidal binaries are not necessarily eclipsing and tend to have very short periods. The sinusoidal nature of the light curves arises from the fact that stars in an extremely close orbit are not spherical: they bulge towards one another as the force of gravity from the companion and the centrifugal force acts on each star. When they are observed side on, the bulges are at their maximum. When the stars are lined up, the bulges will be at their least visible. Where phase is designated by ϕ, the minima in these light curves will occur at $\phi = 0$ and $\phi = 0.5$ (stars are in line), and the maxima at $\phi = 0.25$ and $\phi = 0.75$ (stars are edge on), assuming zero eccentricity.

It is unlikely that semi-detached, contact and ellipsoidal binaries would be mistaken for transiting exoplanets, due to their distinctive light curves, and Coughlin et al. (2016) reported that systems known to have a morphology parameter > 0.6 were excluded from the *Kepler* Data Release 24 (DR24) processing and later iterations. Detached systems, however, may be another matter.

1.5.1 Grazing eclipses

KIC 9027841 and KIC 3113266 in Fig. 1.7 (bottom panel) are both good examples of grazing eclipses in a detached binary system of a type that may be mistaken for transiting exoplanets, without closer examination.

With primary eclipse depths of 900 and 5,000 ppm respectively, both are comparable to planets. For example, if the transit of a planet with the radius of Jupiter could be observed

around a star with the radius of the Sun, the transit depth would be 10,000 ppm; similarly, an Earth-like planet transiting a Sun-like star would have an transit depth of 80 ppm. These two transits fall in between these examples.

One major clue that these are eclipsing binaries rather than close-in planets is the shape of the transit in the light curve. An exoplanet transit is made up of three phases: the ingress, from the point where the the surface of the planet appears to make contact with the surface of the star, to the point where the full disc of the planet is just over the surface of the star; the egress, which is simply a reverse of the ingress and occurs at the end of the transit; and the main part of the transit, when the full disc of the planet is passing over the disc of the star. The ingress and egress are marked by a relatively steep drop/increase in flux respectively, and while the mid part of the transit is not completely flat, due to limb darkening, the general effect is a "flower pot" shape: two sloping sides with a (relatively) flat bottom.

In a grazing eclipsing binary, the full disc of the eclipsing star never covers the star that is being eclipsed. Hence, the ingress and egress dominate. The result is a distinctive 'V' shape, as seen in KIC 3113266 in Fig. 1.7. The 'V' shape in KIC 9027841 is gentler than in KIC 3113266, but, combined with the observable secondary eclipse, it is still distinctive enough for this object to be declared an eclipsing binary by the *Kepler* Eclipsing Binary Catalogue, and a false positive on NExSci.

Care should, however, be taken not to designate all V-shaped transits as eclipsing binaries: a grazing eclipse by a hot Jupiter may also result in a V-shaped light curve.

1.5.2 Eclipsing white dwarfs

White dwarfs (WDs) mark the end stage in the life of most stars. Mass has been lost, particularly in the form of planetary nebulae during the thermally pulsing asymptotic giant branch (TP-AGB) phase, and the stellar remnant is electron degenerate. This means that, although a typical white dwarf may have a mass of 0.59 $M_{\odot}$, its radius is likely to be typical of a terrestrial planet rather than a star. The light curve of an eclipsing binary in which a white dwarf passes behind and in front of its companion would be expected to have a similar shape to that of a planet passing in front of its host star. Follow-up observations are required to distinguish a white dwarf from an exoplanet.

That this is not always easy is shown by the case of KOI-256. Muirhead et al. (2013) described this system as a mutually eclipsing post common envelope binary: that is, stars came into contact and underwent a common envelope phase, as described above, with a main sequence (MS) star orbiting within the envelope of an evolved star, such as a TP-AGB. The TP-AGB star expelled its outer layers as it evolved into a WD, while the MS star lost angular momentum as it orbited within the TP-AGB's atmosphere. The eventual result was

an MS/WD binary in a close orbit: in this case the orbital period is 1.38 days. Muirhead et al. (2013) found that the WD has $M_* = 0.592 \pm 0.089$ $M_\odot$, $R_* = 0.001345 \pm 0.00091$ $R_\odot$, $T_{eff} = 7100 \pm 700$ K, and the MS star has $M_* = 0.51 \pm 0.16$ $M_\odot$, $R_* = 0.540 \pm 0.014$ $R_\odot$, $T_{eff} = 3450 \pm 50$ K.

Sokov et al. (2012) described the same object as a confirmed planet with $R_P = 1.83 \pm 0.16$ R_J with an unusual orbit, based on observations made by ground based telescopes. This is in contrast to the 5.6 $R_\oplus$ found by Muirhead et al. (2013). The community appears to accept Muirhead et al. (2013) rather than Sokov et al. (2012) (see, for example, Mann et al. (2013), Ballard and Johnson (2016) and Gaidos et al. (2016)).

The NASA Exofop website (https://exofop.ipac.caltech.edu/cfop.php) lists this KOI as a false positive, as does NExSci. Exofop holds five spectroscopic observations along with one imaging observation, but no RV observations have been uploaded to the site. Muirhead et al. (2013) made reference to RV measurements, and these would have confirmed that the mass of the smaller body (by radius) was stellar rather than planetary. The *Kepler* False Positive Working group has certified this KOI (*Kepler* Object of Interest) as a false positive with a significant secondary (http://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=fpwg), quoting Muirhead et al. (2013) in the comments field. Interestingly, the radius listed in the official reports, which are publicly available by following the links in the NExSci site, is 5.5 $R_\oplus$ in the Q1-Q16 analysis (report generated 17 August 2013), 6.4 $R_\oplus$ in the DR24 analysis (report generated 20 September 2014) and 2.0 $R_\oplus$ in the DR25 analysis (report generated 30 January 2016). None of these reports identify this KOI as an eclipsing binary, so the follow up work by Muirhead et al. (2013) was essential in identifying this system as an eclipsing MS/WD binary.

Another difference between a WD and a planet is the effective temperature. Muirhead et al. (2013) found the compact object to be significantly hotter than the MS star, which is consistent with it being a WD. A planet would be much cooler. Where a secondary eclipse can be observed (ie when the larger body passes between the observer and the smaller body), as is the case here, the *Kepler* pipeline estimates the temperature of the transiting body based on observations and using four assumptions which also assume that the transiting object is a planet[4]:

1. The incident stellar flux and the radiated heat from the "planet" are in thermodynamic equilibrium.

2. The "planet" has a Bond albedo of 0.3, which means that 30% of the incident stellar flux is re-radiated and 70% is absorbed.

3. Both the "planet" and the star are black bodies.

4. Heat is evenly distributed between the day and night sides of the "planet".

[4]https://exoplanetarchive.ipac.caltech.edu/docs/API_kepcandidate_columns.html

This estimate can be compared to the predicted maximum effective temperature of a planet at the same distance from the star. Exofop gives an estimated planet temperature of 734 K, consistent with the working assumption that the transiting object is a planet, again emphasising the importance of the follow-up work by Muirhead et al. (2013) in identifying the effective temperature of the WD as T_{eff} = 7100 $\pm$ 700 K.

The conclusion from this test case is that automated data vetting alone cannot separate a white dwarf from a genuine planet: rigorous follow-up work is required.

1.5.3 Blended eclipsing binaries

Although we observe space as a two dimensional surface, it is in fact a three dimensional volume. The classic analogy is that stars that we put into constellations based on two dimensional shape, such as the W of Cassiopeia, are often not related to one another at all. It is essential to bear the three dimensional nature of space in mind when analysing the results of a transiting exoplanet survey. Where a target is the star you intend to observe, there may be a faint eclipsing binary in the vicinity of the target, whose position is indistinguishable from that of the target in the initial survey due to the plate scale, and which can only be identified as a false positive after follow-up observations.

One initial check performed in surveys such as *Kepler* and PLATO is to test for ephemeris matches: is there another transit with exactly the same orbital period, at exactly the same time, elsewhere within the field of view? In such situations, the actual eclipsing binary or variable star is referred to as the "parent" and the contaminated targets are described as the "children". This test, however, will not reveal all false positives: only those where the parent is far enough from the child to be observed independently.

Blending of eclipsing binaries occurs when the flux from the target star is combined with the flux from an eclipsing binary, for example when the PSF of the background star is fully within the PSF of the target or by having an overlapping PSF with the target. In the case of *Kepler*, charge transfer along columns is known to cause ephemeris matches to distant stars.

Where the flux of two or more systems is combined, the observed depth of any transit will be less than the actual depth (it is 'diluted'), so a transit that would otherwise appear clearly stellar from Eq 1.1 may now appear planetary.

The normalised transit depth, δF, can be found by comparing the in-transit and out of transit flux, as in Eq 1.10. The standard equation for finding the difference between the apparent magnitudes of two stars when the flux of the two stars is known is

$$m_1 - m_2 = 2.5 log_{10} \left(\frac{F_2}{F_1} \right) \tag{1.13}$$

By substituting the observed transit depth δF_{obs} for F_1, the actual transit depth δF_{act} for F_2, and Δm, the difference in apparent magnitude between the actual eclipsing binary and all stars contributing to the observed flux (including both the eclipsing binary and the target), for $m_1 - m_2$, and rearranging, the observed depth can be found from the actual depth:

$$\delta F_{obs} = \frac{\delta F_{act}}{10^{\Delta m/2.5}}. \tag{1.14}$$

and Eq 1.14 can be easily rearranged to find the actual depth from the observed depth, where Δm is known. If there is only one system contributing to the observed flux, Eq 1.14 will return an observed depth that is equal to the actual depth as Δm would be zero: if there is more than one system, a diluted observed depth will be reported. In the context of PLATO, with large pixels of 15", Eq 1.14 has the advantage of allowing *all* background flux contaminating the PSF of the target star to be included, thereby allowing dilution due to blending with non-eclipsing background stars to be taken into account in the case where the target is actually an eclipsing binary, as well as estimating dilution of background eclipsing binaries.

Whether a given transit is due to an exoplanet or a background eclipsing binary is a question that the *Kepler* science team has investigated, as well as the issues that arise from plate scale. Morton and Johnson (2011) define an "exclusion radius" r within which it is not possible from *Kepler* data alone to tell if a transit is on the target or due to a background eclipsing binary:

$$r = 1.17'' \sqrt{10^{-0.4(11-m_{Kp})}} \left(\frac{\delta F}{1.5 \times 10^{-4}} \right)^{-1} \tag{1.15}$$

where m_{Kp} is apparent *Kepler* magnitude. The constant of 1.17" is based on the analysis performed on *Kepler*-10 b in Batalha et al. (2011). Using Eq 1.15 an appropriate exclusion radius can be calculated for any target, as long as m_{Kp} and δF, both observables, are known.

Fig. 1.8 shows some of the diagnostic techniques used by the *Kepler* science team to rule out a background eclipsing binary scenario in the case of the second planet in the *Kepler*-10 system, *Kepler*-10 c. This is included as an example of the procedures available to the *Kepler* science team, which may or may not be applicable to the PLATO science team. The information for this plot came from the DR25 report, which is publicly available through NExSci. Where *Kepler*-10 b has a period of 0.84 days, *Kepler*-10 c has a period of 45.3 days, so there are only two transits per quarter, at most, to assess. The whitened folded average time series reveals an eclipse depth of 477.1 $\pm$3.7 ppm, with a shape that is consistent with a planetary transit. The pixel level diagnostics show that the source of the difference in flux appears to be the target star, *Kepler*-10: the Q9 difference plot has been randomly selected for inclusion in Fig. 1.8, but each quarter is assessed independently. Combining the difference information from all quarters, an average position for the transit source is located, with 1-σ

and 3-σ uncertainties indicated by a cross and circle respectively. The target falls outside the 1-σ range but inside the 3-σ circle in the case of *Kepler*-10 c. However, superimposing this information on an image from the United Kingdom InfraRed Telescope (UKIRT) shows that the source of the transit is clearly on the target star, and even allows fainter background stars to be visually identified. Note the difference in orientation between the pixel scale plot, which refers to the orientation of *Kepler* at the time the observations were made and where the axes are labeled by CCD column and CCD row, and the plots in which the centroid is indicated, where the axes are right ascension (RA) and declination (dec). Taking all this information together, it would appear that *Kepler*-10 c is indeed a planet orbiting the observed target, and not a background eclipsing binary.

As mentioned previously, Hełminiak et al. (2017) found KIC 4150611 (KOI 3156) could be resolved spectroscopically to three point sources, one of which, labelled C, was most likely to be a background eclipsing binary. Since the source was observed as one target, the eclipses in A and B, as well as C, would be blended, in the manner described in Eq 1.14. The PLATO science team will have access to the full Gaia data, so it is probable that background stars in the vicinity of an observational target would be known and can be accounted for. However, care will still be required.

Blended eclipsing binaries will have an apparent eclipse depth shallower than their actual eclipse depth. Blending may also occur if a target believed to be a single target is in fact an unresolved triple or other hierarchy. And the transit of a planet in orbit around one of the stars in an unresolved binary system will also be diluted by blending, and hence the planet will appear to have a smaller radius than the true value.

1.6 Population synthesis as a tool to understand the Galaxy

The previous sections in this chapter have dealt with observations of transiting exoplanets and of eclipsing binaries, and how we distinguish them using existing instruments. Each instrument is unique, and to understand the false positive rate due to eclipsing binaries in PLATO it is important to build a model Galaxy that can be used to test the instrument's expected performance.

Modelling is an important tool in understanding a complex system. Population synthesis is the class of model that is used to test our understanding of the Galaxy. Generally held assumptions include:

1. The Galaxy is spiral and has a central bar.
2. The Galaxy consists of a central bulge, a disc and a halo.
3. Star formation occurs close to the Galactic plane.

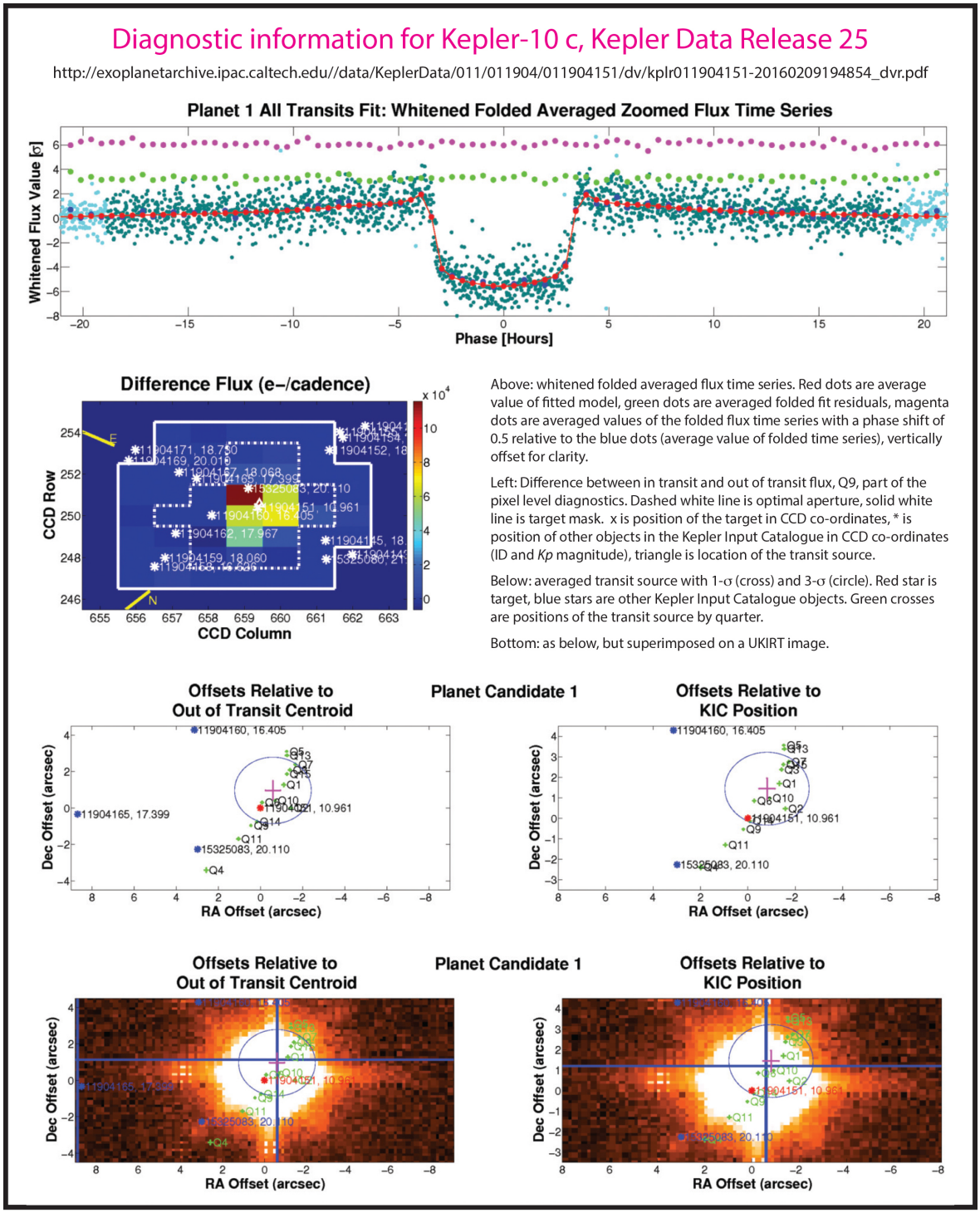


Fig. 1.8 Selected diagnostic information from *Kepler* DR 25, *Kepler*-10 c, in particular relating to position. Data publicly available on the NExSci website.

4. The disc has two distinct populations, the thick disc, consisting of low metallicity stars and the thin disc, consisting of higher metallicity stars.

Since we are in the Galaxy, we cannot observe it from the outside. However, the assumption of a spiral structure appears sound. Observations of other galaxies show that star formation occurs in spiral galaxies rather than elliptical galaxies, and star formation is observed in the Galaxy, for example in the Pleiades. Observations also indicate that the typical structure of a spiral galaxy is a central bulge, which contains a supermassive black hole, along with two or more spiral arms in which star formation takes place. Mapping by NASA's Wide-field Infrared Survey Explorer (WISE)[5] (Camargo et al., 2015) (Fig. 1.9) indicates that the Galaxy's central bulge is a bar, with several arms radiating from each end of the bar. The innermost arm in each case rejoins the other end of the bar, to form a roughly oval shape; the other arms move out to form spirals. The Sun is located in the Orion Spur of the Sagittarius Arm, a relatively unimportant arm that emerges at the far end of the central bulge, as observed from Earth. Most dust and molecular clouds are located close to the Galactic plane, so the Galactic plane is a logical place for star formation to occur.

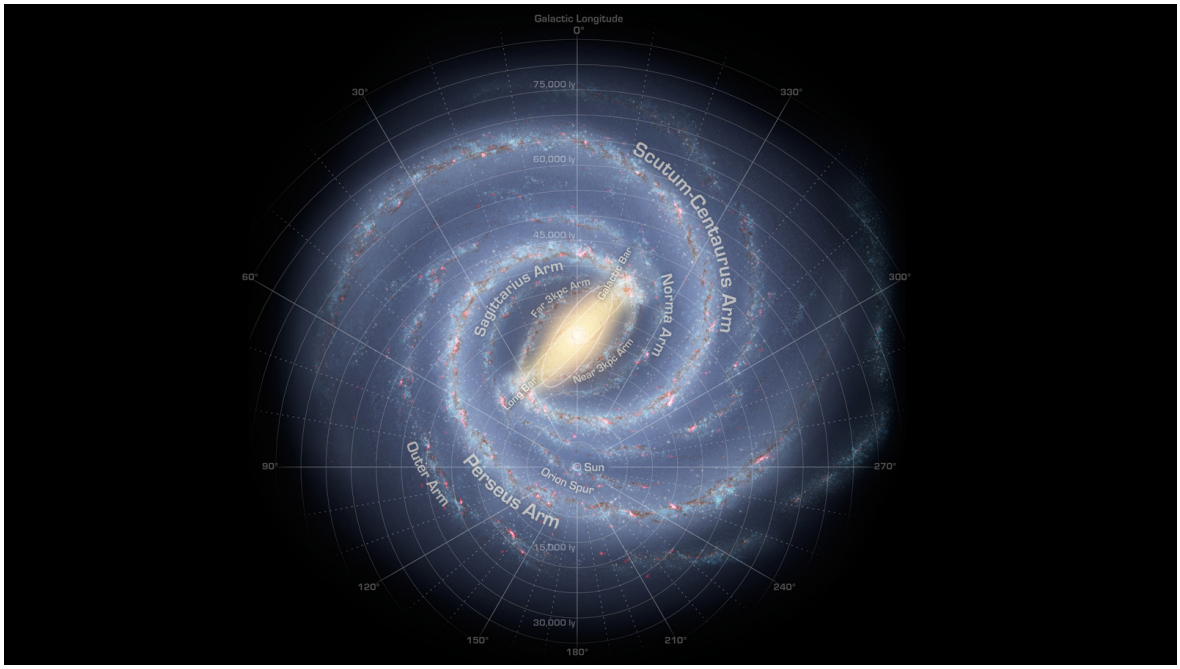


Fig. 1.9 Schematic diagram of the Galaxy from data from NASA's Wide-field Infrared Survey Telescope (WISE), downloaded from https://www.jpl.nasa.gov/news/news.php?feature=4612

Jurić et al. (2008) used observations from the Sloan Digital Sky Survey (SDSS) to confirm the model of bulge, thick and thin discs, and halo. Studying the number density of M class MS stars within 2 kpc of the Sun, the authors find that the thin disc has a scale height of 300

[5]https://www.jpl.nasa.gov/news/news.php?feature=4612

pc and scale length of 2600 pc, and the thick disc, populated by older stars, has a scale height of 900 pc and a scale length of 3600 pc. Jurić et al. (2008) found that the halo is likely to be oblate rather than spherical, with a best-fit axis ratio c/a of 0.64.

With this understanding of the basic structure of the Galaxy, population synthesis models can be derived. Two of the best known are TRILEGAL and the Besançon model, with the Gaia Universe Model Snapshot (GUMS) and Galaxia being derived from the latter. In addition, the Open University has over the past two decades developed its own model, the Binary Stellar Evolution Population Synthesis model (BiSEPS), which will be discussed in Chapter 2.

1.6.1 TRILEGAL

The TRILEGAL population synthesis code, the TRIdimensional modeL of thE GALaxy, is first described in Girardi et al. (2005). Updates are discussed in Girardi et al. (2012). It is available for public use through an online interface[6], last updated on 23 October 2017, and has a time limit of 10 minutes CPU time per request.

Girardi et al. (2005) stated that the code, which is written in C, extracts stars of a given mass, age and metallicity from a database of stellar tracks, and seeds them into a model Galaxy. The code is tuned to the specifications of the instrument which could make the observations that are being simulated, including the passband. The "perfect photometric catalogue" produced can be later degraded as required to allow for noise and incompleteness. Girardi et al. (2005) described the five inputs into TRILEGAL:

1. Stellar evolutionary tracks from zero age main sequence (ZAMS) to 25 Ga or the end of the TP-AGB stage, whichever comes first. These tracks are described in Girardi et al. (2000). Protostars and compact remnants are not simulated.

2. Bolometric corrections as described in Girardi et al. (2002) and absorption coefficients as described in Girardi et al. (2004). The bolometric corrections allow for the fact that any passband will only see part of a star's output, while the absorption coefficients allow for extinction by interstellar dust.

3. The initial mass function (IMF), which defines the proportion of less massive to more massive stars. The default for TRILEGAL is the log-normal IMF is described in Chabrier (2001), but other options are available, including the exponential IMF described in Larson (1986) and the segmented power laws described in Salpeter (1955) and Kroupa (2001).

4. Star formation rate (SFR) and the age metallicity relation (AMR) for the different Galactic components. The thick disc has a constant SFR and AMR for the period 11-12 Ga. The halo is set at an age of 12-13 Ga and offers two options for AMR. For the thin disc, there

[6]http://stev.oapd.inaf.it/cgi-bin/trilegal

is the option of a one-step or two-step SFR with the AMR taken from Fuhrmann (1998) with α enhancement. The AMR for the bulge is based on Zoccali et al. (2003).

5. Geometry of the Galaxy components. The thin and thick discs can be described either by a double exponential or by an exponetial in the radial direction and a squared hyperbolic secant in height. The halo is described as an $r^{1/4}$ spheroid, with adjustable oblateness. The bulge is modeled as a triaxial truncated spheroid.

Girardi et al. (2005) went on to describe the calibration of `TRILEGAL`, and its use in simulating the output of various surveys, including 2MASS and Hipparcos.

While the public interface is ideal for small scale studies, in particular regions of sky that can be simulated within the time limit of 10 minutes CPU time, it is not suitable for simulating large regions of sky, as required in the PLATO simulations.

1.6.2 The Besançon model, `GUMS` and `Galaxia`

The Besançon model, described in Robin et al. (2003) has been used to create the Gaia Universe Model Snapshot (`GUMS`) (Robin et al., 2012), which can be publicly accessed through the VizieR service[7]. `Galaxia` is a tool which samples stars from the Besançon model and has been used to generate a mock Gaia DR2 catalogue which can also be publicly accessed[8] (Rybizki et al., 2018). The underlying Besançon model can also be publicly accessed, coming back on line on 17 April 2017 after a cluster upgrade.[9].

Robin et al. (2003) described the assumptions behind the Besançon model. It is assumed that star formation is constant in the thin disc but occurred in one burst in the thick disc (11 Ga), halo (14 Ga) and bulge (10 Ga). The thick disc, halo and bulge are modelled to each have a constant metallicity $[Fe/H]$ (dex) of -0.78 ± 0.30, -1.78 ± 0.50 and 0.00 ± 0.40 respectively, while metallicity in the thin disc evolves with age: 7-10 Ga -0.37 ± 0.20, 5-7 Ga -0.14 ± 0.17, 3-5 Ga -0.07 ± 0.18, 2-3 Ga 0.01 ± 0.11, 1-2 Ga 0.03 ± 0.10, 0.15-1 Ga 0.03 ± 0.12, 0-0.15 Ga 0.01 ± 0.12. The thick disc, halo and bulge have one IMF per component, ie a continuous slope, while in the thin disc there is a break at 1 $M_{\odot}$. Disc kinematics are included in the calculations in three dimensions, and it is assumed stars drift away from the Galactic plane as they age, as well as orbit the Galactic bulge.

The original Besançon model does not account for interstellar clouds or spiral arms, a feature it has in common with `TRILEGAL`. `GUMS`, created from the Besançon model, incorporates two spiral arms. Robin et al. (2012) gave the parameters of these arms as having an internal radius of 3.426 kpc each, a pitch angle of 4.027 and 3.426 radians respectively and a phase angle of start in radius of 0.188 and 2.677 radians respectively. The amplitude and thickness

[7]http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=VI/137
[8]http://dc.g-vo.org/tableinfo/gdr2mock.main
[9]http://model.obs-besancon.fr

in the plane of the arms are also defined as of 1.823 and 4.804 for the first arm and 2.013 and 4.964 for the second arm respectively, but without units of measurement: presumably these are kpc.

The mock Gaia DR2 catalogue was generated with Galactic warp switched on and sets the solar zero-point to $(X,Y,Z) = (-8.0, 0.0, 0.15)$ kpc with velocities $(U,V,W) = (11.1, 23.08, 7.25)$ km s^{-}1, and utilises updated extinction models. Its stated limitations include excluding stellar binaries and stellar remnants, as well as extragalactic sources, as these classes of objects are not included in Gaia DR2.

GUMS was generated with the aim of predicting the number of objects that Gaia would observe, in order to constrain the amount of data that would be likely to be downloaded so that sufficient computing facilities would be available to process it. It is a snapshot of the Galaxy at one moment in time, hence the name. It is a sophisticated model, based on a Galaxy model matched as closely to observations as could be achieved. As with TRILEGAL the online interface is useful for checking assumptions about discrete areas of the sky, but not for the large scale simulations required for PLATO.

The mock-Gaia DR2 catalogue was published shortly before the release of Gaia DR2. It is intended to provide the opportunity to practice writing queries ahead of the release of DR2, thereby saving researchers time when the data is available. Obviously, for actual queries researchers will refer to Gaia DR2 itself.

Online access to the Besançon model offers a choice of two photometric systems, Johnson-Coussins or CFHT-Megacam; output in the form of a catalogue or a table with differential counts; and with or without kinematics. The online form covers a wide range of observational parameters, and an email is sent when results are available. A 1° x 1° field centred on $l = 65°$, $b = 30°$ completed in seconds (20 April 2017), a larger simulation of $55° < l < 75°$, $20° < b < 40°$ required five minutes. In both simulations, the position of every star was $l = 65°$, $b = 30°$, and there was no way to distinguish binaries and higher hierarchies from single stars.

1.7 Summary: Chapter 1

This chapter has explored the potential of transiting exoplanet surveys in estimating the radii of exoplanets which, when a mass estimate is also available from other observations, is essential in constraining the density of individual planets and hence the bulk characteristics of the exoplanet population. PLATO, a planned transiting exoplanet survey mission planned for launch in 2026, has the potential to greatly enhance our understanding of the transiting exoplanet population.

However, certain classes of eclipsing binary can produce false positives in such surveys: in particular binaries in which one component is a white dwarf, grazing eclipsing binaries and background blended eclipsing binaries.

This project combines data from the most prolific transiting exoplanet survey to date, the *Kepler* main mission, with stellar population synthesis to estimate the risk of contamination from eclipsing binaries in the PLATO field. Next I describe the population synthesis code used in this work, `BiSEPS`.

Chapter 2

BiSEPS

BiSEPS, the Binary Stellar Evolution Population Synthesis code, is the main numerical tool utilised in this thesis. This chapter summarises the main features of BiSEPS, its advantages and limitations, and describes the calibration checks carried out for the simulations used in this project.

2.1 History of BiSEPS

2.1.1 Binary Stellar Evolution

BiSEPS was originally developed at the Open University by Dr Bart Willems, under the supervision of Dr Ulrich Kolb. Using fitting formulae from Hurley et al. (2000) and Hurley et al. (2002), it describes the evolution of both single stars and binaries from ZAMS to compact remnant. Further developments were made to the code used here by Dr Rob Farmer, also under the supervision of Ulrich Kolb.

Hurley et al. (2000) examined single star evolution, and designated 16 evolutionary classes, described in Table 2.1, providing simple fitting formulae for each class to describe its evolution as a function of mass, age and metallicity. A schematic guide to possible evolutionary paths through these classes was published in Hurley et al. (2000), their Fig. 19, and has been adapted here as Fig. 2.1. The single star evolutionary code associated with Hurley et al. (2000) was incorporated by Willems in the earliest incarnation of BiSEPS.

Some of the evolutionary paths shown in Fig. 2.1 are only possible in a binary system. Examples include mass gain by a carbon/oxygen white dwarf (CO WD), which on exceeding the Chandrasekhar limit of 1.4 $M_{\odot}$ will explode, leaving what Hurley et al. (2000) described as a "massless remnant": this is most likely to occur through accretion in a binary system, where the less evolved star fills its Roche lobe and donates mass to the CO WD. Another example is the formation of helium white dwarfs (He WD). Low mass stars evolve more

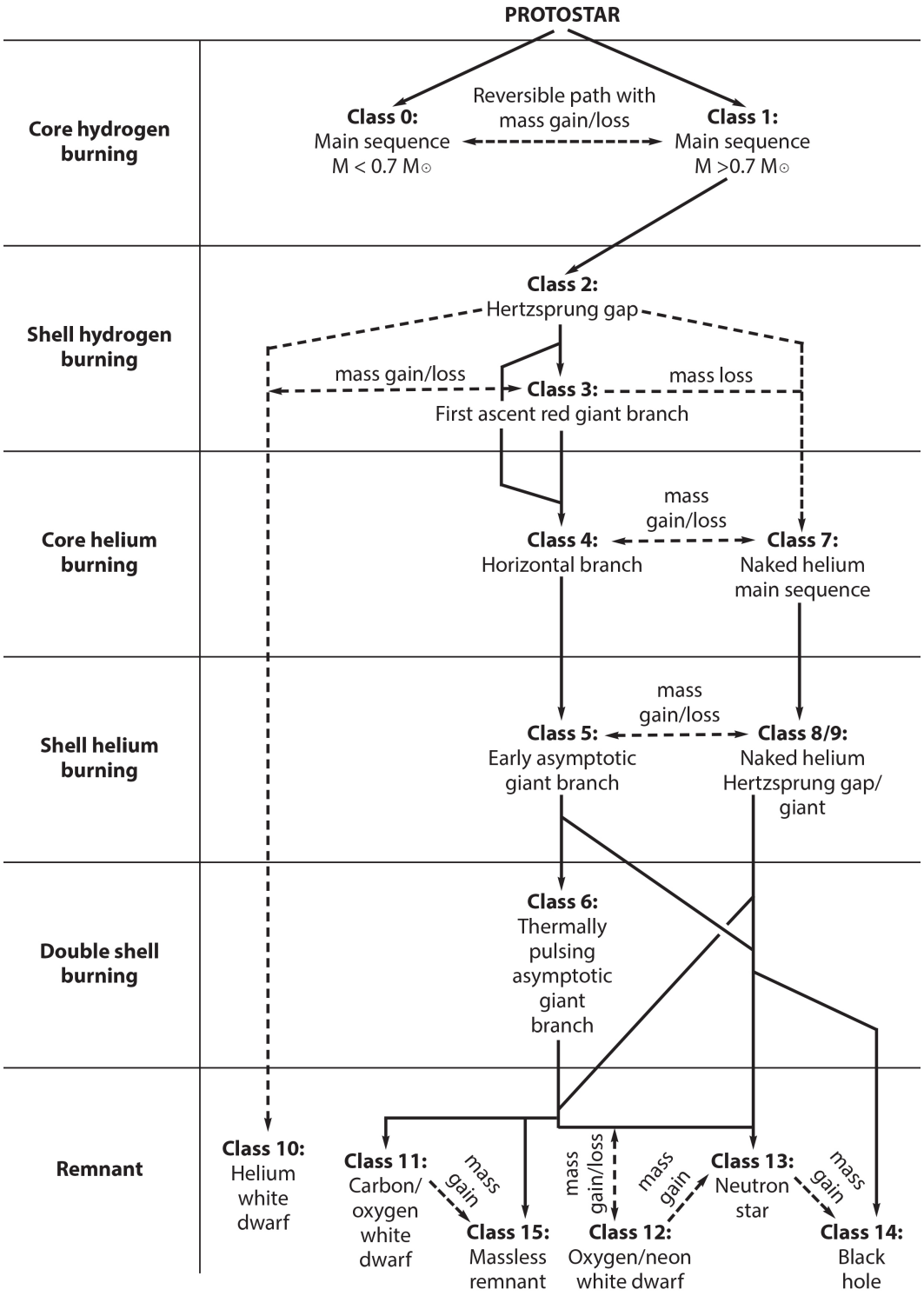


Fig. 2.1 Stellar evolutionary paths and classes as descrbed in in Hurley et al. (2000) and adapted from their Fig. 19.

Table 2.1 BiSEPS stellar classes, from Hurley et al. (2000)

Class Number	Description
0	MS star $M \leq 0.7$ M$_\odot$ deeply or fully convective
1	MS star $M > 0.7$ M$_\odot$
2	Hertzsprung gap (HG)
3	First ascent red giant branch (RGB)
4	Core helium burning (CHeB)
5	Early asymptotic giant branch (E-AGB)
6	Thermally pulsing asymptotic giant branch (TP-AGB)
7	Naked helium star, main sequence
8	Naked helium star, Hertzsprung gap
9	Naked helium star, giant branch
10	Helium white dwarf (He WD)
11	Carbon/Oxygen white dwarf (CO WD)
12	Oxygen/Neon white dwarf (ONe WD)
13	Neutron star
14	Black hole
15	Massless remnant

slowly than high mass stars and, for example, any single star < 10 Ga in age and < 0.9 M$_\odot$ in mass will still be on the main sequence at the current time epoch. A He WD forms when conditions in the core are such that helium ignition cannot occur, which in a single star implies a mass at ZAMS of < 0.5 M$_\odot$. Given the age of the Universe at $\approx$ 13.6 Ga and hence the upper limit of the age of the Galaxy, He WDs cannot have formed from single star evolution. The fact that they are observed in the Galaxy means they must have formed in a close binary system, where mass transfer has reduced the mass of the donor to the point where core helium ignition is no longer possible in the initially more massive star.

Binary evolution was considered in Hurley et al. (2002) and, working independently, Willems had the advantage of early sight of this paper, so was able to incorporate the concepts into `BiSEPS`. Features of binary evolution considered in Hurley et al. (2002) include wind accretion, orbital changes due to mass variations, tidal evolution, gravitational radiation and magnetic braking, supernovae kicks, Roche lobe overflow and common envelopes, coalescence and collisions. Some of these features, such as wind accretion and Roche lobe overflow, are incorporated into `BiSEPS`: others have yet to be included, including coalescence and collisions.

`BiSEPS` produces an evolutionary sequence of models for a binary system with components of a given initial mass. The mass, radius, luminosity and effective temperature of both components is recorded for each model, along with orbital period, metallicity, the age at

the start and end of each model in the sequence, and a series of keywords to indicate the evolutionary status of the two components, whether mass transfer has ever occurred, if mass transfer is currently occurring and if so, of what kind. The models are saved in a look-up table. Single stars are generated by creating a secondary of < 0.1 $M_\odot$ and placing it at a considerable distance from the primary, with an orbital period of $> 10^9$ days. The primary in such a system evolves as a single star. The secondary in such a "single" system is faint by design, and does not contribute to the magnitude of the system as a whole, except in the case of a very low mass primary. Very low mass single stars are unlikely to contribute to synthetic Galaxy models, however, except as background stars. With a luminosity in each component of $\approx 10^{-3}$ $L_\odot$ in the extreme case of a 0.1 $M_\odot$ primary (which will have an equal mass secondary), it is unlikely that a very low mass "single" star will exceed the magnitude limit used in an on-sky survey. Single stars are also saved in a look-up table, with the same information as for binaries.

Evolution ends when one of the following conditions is met:

1. The age at the start of a model within an evolutionary sequence is 15 Ga.
2. A common envelope is formed resulting in coalescence.
3. A black hole is formed.
4. A massless remnant is formed.

Examples of the power of `BiSEPS` in describing binary systems include studies of wide binary millisecond pulsars (Willems and Kolb, 2002), pre-low-mass X-ray binaries (Willems and Kolb, 2003), detached WD-MS binaries (Willems and Kolb, 2004), deriving accurate parameters for the β Cephei star λ Scorpii (Uytterhoeven et al., 2004), the effects of circumbinary discs on angular momentum losses and the orbital period distribution of CVs below the period gap (Willems et al., 2005), (Willems et al., 2007), the role of magnetic braking in CV systems crossing the period gap (Davis et al., 2008) and post common envelope binaries (Davis et al., 2010).

2.1.2 Population Synthesis

Using an understanding of the structure of the Galaxy, the models generated in the stellar evolution part of BiSEPS can be seeded into a synthetic Galaxy. The work for this was started by Bart Willems and extended by Rob Farmer. Synthetic star fields have been used in a study of eclipsing binaries in the SuperWASP survey (Willems et al., 2006), a study of the true stellar parameters of the initial *Kepler* target list (Farmer et al., 2013) and a study of asteroseismic binaries in *Kepler* data (Miglio et al., 2014).

The parameters describing the Galaxy used in this project are those adopted by Willems and Farmer, and are described in more detail in Farmer et al. (2013) and references therein.

The thin and thick discs are simulated, but not the halo or the bulge. Two metallicities are available, one for each disc, with $Z_\odot$ ($Z = 0.02$) assigned to the thin disc and 0.165 $Z_\odot$ ($Z = 0.0033$) assigned to the thick disc. The reasons behind the choice of these metallicities, based on the work of Haywood (2001) and Gilmore et al. (1995), are discussed in more detail in Section 2.2.1, subhead "Metallicities." Other metallicities can be simulated with the stellar evolution code, but with the population code limited to two metallicities, these have been selected as the best match to observations. In `BiSEPS`, it is assumed that star formation occurred in the thick disk 13-10 Ga ago, and that star formation in the thin disc occurred during the past 10 Ga only. Both discs are modelled with continuous star formation, at a rate of 1.2 star yr^{-1} with > 0.8 $M_\odot$: this value was obtained by Farmer et al. (2013) when calibrating a synthetic *Kepler* field to the initial *Kepler* input catalogue. A Kroupa IMF with two breaks is used (Kroupa, 2001), with a slope of -1.23 for 0.1-0.5 $M_\odot$, -2.2 for 0.5-1.0 $M_\odot$ and -2.7 for > 1.0 $M_\odot$. Drimmel extinction is used (Drimmel et al., 2003).

The position of the Sun is set at 8.5 kpc from the centre of the Galaxy and 0.030 kpc above the Galactic plane. The scale length of the thin disc is 2.8 kpc (radial, h_R) and the scale height is 0.3 kpc (vertical, h_z), whereas the scale length of the thick disc is 3.7 kpc (h_R) and the scale height is 1.0 kpc (h_z). These scale heights and lengths are based on the study by Jurić et al. (2008) described in Section 1.6. Both discs are modelled as a double exponential of the form

$$\Omega(R,z) = n_0 e^{(-R/h_R)} e^{(-|z|/h_z)} \tag{2.1}$$

where

$$n_0 = \frac{1}{4\pi h_R^2 h_z} \tag{2.2}$$

The binary fraction is set at 0.5: ie half the initial population will be binary systems, and half will be single systems. As discussed in (Farmer et al., 2013), this is essentially arbitrary. Marks and Kroupa (2012), while considering the binary population of young clusters, demonstrate from simulations that it is possible to start with a formal binary fraction of unity and for the cluster to evolve over time to a lower binary fraction. They note in their introduction that younger clusters appear to have a higher binary fraction than older clusters with a similar stellar density. Reproducing such a trend over time is currently beyond the scope of `BiSEPS`.

More important is addressing the fact that the binary fraction does not accurately represent the distribution of observed binaries, where high mass stars are more likely to be binary and low mass stars are more likely to be single (Fischer and Marcy (1992), Lada (2006)), and which is recommended for future work (Chapter 9) to quantify the effect on my results. High mass (O-B class) binaries are more likely to be observable than low mass (M-class) binaries

and are more likely to exceed the magnitude limit in a synthetic population, so higher mass background eclipsing binaries are more likely to contaminate a transiting exoplanet survey. However, the initial mass function indicates that O-B class stars are less common than M class stars. These competing factors need to be compared quantitatively in future work.

The initial mass ratio (IMR) is set as flat, as is the initial period distribution: so there is an equal probability of simulating a binary system with any given mass ratio and any given orbital period. Comparing the results of the simulated output with observations, as will be described in Chapter 7, provides a deeper understanding of the initial mass ratio and initial period distribution of short period binaries in the observed Galaxy.

The population synthesis code runs through the models provided by the stellar evolution look up tables, seeding an appropriate number of stars in a given on-sky area, defined by l and b co-ordinates, taking into account distance d, extinction, initial mass function and the magnitude limit set by the user. A wide variety of passbands are available, including *Kepler*, Sloan, Johnson-Cousins-Glass, Stroemgren, 2MASS, Ogle, Spitzer and UKIDSS.

Input parameters using heliocentric co-ordinates (l, b, d) are converted to Galactocentric co-ordinates (R, z) using

$$R = \sqrt{(d^2 \cos^2 b - 2dR_{\odot} \cos b \cos l + R_{\odot}^2)} \tag{2.3}$$

and

$$z = d \sin b + z_{\odot} \tag{2.4}$$

In addition to the information from the stellar evolution code, the population synthesis code records the on-sky location (l, b), distance (d) and extinction of the randomly seeded systems, absolute and apparent magnitudes of the system in the *Kepler* (Kp) passband, and apparent magnitudes in selected passbands. For this project, the selected passbands are Kp, D51, Sloan u', g', r', i' and z', and Johnson-Cousins-Glass J, H and K. The magnitude limit is set at $Kp \leq 26$, to enable a deep, unresolvable background to be simulated.

2.2 Advantages and limitations of BiSEPS

2.2.1 Limitations

There are a number of limitations associated with BiSEPS but, as will be shown in Section 2.3, these do not affect its ability to synthesise a Galaxy which matches current observations.

Eccentricity

BiSEPS assumes zero eccentricity. While binary systems will circularise over time, when $P > 10$ days a wide range of eccentricities is possible. The depth and duration of each transit (primary and secondary) in an eclipsing binary will depend on the relationship between the observer's line of sight, periastron and apastron, as well as the angle of inclination, i.

Assuming that the primary transit in an eclipsing binary occurs at $\phi = 0$, when $e = 0$ the secondary transit will occur at $\phi = 0.5$: but when $e \neq 0$, the secondary eclipse may occur at a different value of ϕ.

Binary systems with $P < 10$ days have generally circularised over the lifetime of the Galaxy, so for such systems the assumption that $e = 0$ holds (Duquennoy and Mayor (1991), Moe and Di Stefano (2017)). For this reason, the work described in Chapter 7 to find appropriate weightings in terms of mass ratio q and orbital period P, using the *Kepler* Eclipsing Binary Catalogue as a calibrator, was carried out exclusively with systems with $P < 10$ days.

Eccentric orbits may alter the binary evolution, as mass transfer may set in much earlier than for circular orbits.

From Kepler's third law, eccentricity will also affect transit duration. Transit duration will be shorter if an eclipse is observed when the stars are relatively close together and longer if an eclipse is observed when the stars are relatively far apart.

Including eccentricity would improve our models. However, including eccentricity in `BiSEPS` will greatly increase the code's run time and is at this stage beyond the scope of this work. It is also not the highest priority for future work.

Further discussion of the issues raised by eccentricity is included in Section 7.2 and 7.3.

Metallicity

The fitting formulae from Hurley et al. (2000) can be applied over a wide range of metallicities. A variety of metallicities is observed in the Galaxy, with each generation of stars becoming more metal-rich than its predecessors: a natural consequence of stellar nucleosynthesis and the seeding of the interstellar medium both through the ejection of planetary nebulae and during supernovae. Metallicity is negatively correlated to effective temperature and luminosity, and positively correlated to stellar radius and age at terminal age main sequence (TAMS) (Hurley et al. (2000), Fig. 3 & 4).

The binary stellar evolution tables of `BiSEPS` could be generated with any metallicity: but the population synthesis part of the code requires one metallicity to be selected for the thin disc and another for the thick disc to keep the computational effort within reason.

Thin disc: Haywood (2001), in a study of metallicity in the solar neighbourhood, found that about half the nearby stars have [Fe/H] > 0. The author found that < 4% of stars in the solar neighbourhood have [Fe/H] < −0.5, and stated that this is consistent with estimates of the thick disc population, which mingles with the thin disc and therefore can be disregarded when finding the average metallicity of the thin disc. Following Haywood (2001), solar metallicity, $Z = 0.02$, is adopted for the thin disc in the `BiSEPS` Galaxy model.

Thick disc: Gilmore et al. (1995) studied the iron abundance of stars 0.5-3.0 kpc above the Galactic plane. They found a broad distribution within the thick disc, from −1.5 < [Fe/H] < 0.0, with a peak ≈ [Fe/H] = −0.8 to −0.7. This corresponds to the value of $Z = 0.0033$ selected to describe the thick disc metallicity.

The metallicity in this work is the same as was used in Farmer et al. (2013), enabling the field generated in that study to be used to calibrate the intrinsic exoplanet population (Chapter 6) and the eclipsing binary population (Chapter 7) against *Kepler* observations before applying the appropriate weightings to the PLATO simulations.

Fig. 2.2 demonstrates that, according to data from the *Kepler* Input Catalogue held on MAST and NExSci (both accessed 29 August 2018), the *Kepler* field is believed to be slightly sub-solar in composition.

However, the difference in metallicity between observations and simulations is likely to have a negligble effect on my results. Stellar tracks were obtained from the online service EZ-Web[1] (a forerunner of the MESA stellar evolution code) for solar mass stars at a metallicity of Z = 0.02 (solar) and Z = 0.01 (equivalent to 0.0 and -0.3 in the measure used in Fig. 2.2). This code was used as it is distinct from `BiSEPS` so would produce independent results, and it is a full stellar evolutionary code rather than fitting formulae. In addition to information on internal processes, such as the central hydrogen mass fraction and the radius of the helium core, data on observables is provided. Table 2.2 presents, as an example, the data on mass, radius, luminosity and surface temperature for a solar mass star of an age of 4.6 Gyr. While there is a clear difference in luminosity and surface temperature, the stellar radii appear comparable at a similar evolutionary stage, especially when the precision of current observational methods are taken into account. Planet radius is assessed against stellar radius. While it would be interesting to examine a synthetic *Kepler* field with a lower metallicity than Z = 0.02 for the sake of completeness, the effect on the results reported in this thesis is unlikely to be significant.

[1] http://www.astro.wisc.edu/~townsend/static.php?ref=ez-web

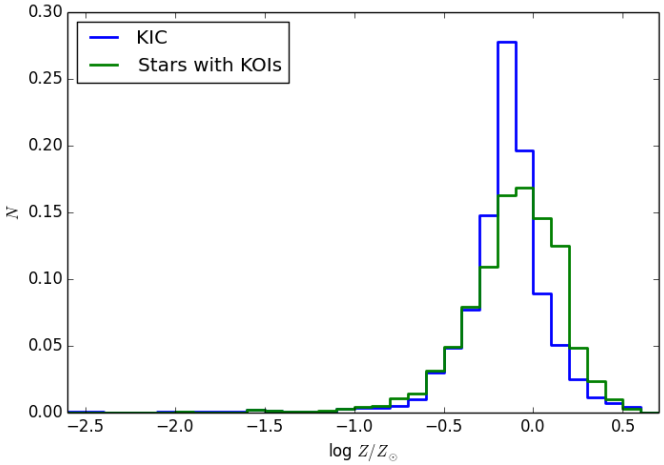


Fig. 2.2 Metallicities of stars in the *Kepler* main mission field, as recorded in the *Kepler* Input Catalogue on MAST (blue line) and for those stars designated KOIs (*Kepler* objects of Interest). Both samples have been normalised to one to highlight the differences in distribution between them. In the *Kepler* catalogues, solar metallicity is set at 0.0. From this data, field stars and KOIs both tend to be slightly sub-solar, but KOI stars tend to have a higher metallicity than the field.

Table 2.2 Effect of metallicity on stellar evolution, solar mass star, at ZAMS, age 4.6 Gyr and TAMS (8.5 Gyr for Z = 0.01, 10.3 Gyr for Z = 0.02). Stellar tracks from the online EZ-Web service.

Parameter	Z = 0.01 ZAMS	Z = 0.02 ZAMS	Z = 0.01 4.6 Gyr	Z = 0.02 4.6 Gyr	Z = 0.01 TAMS	Z = 0.02 TAMS
Radius	0.91 $R_\odot$	0.89 $R_\odot$	1.09 $R_\odot$	1.01 $R_\odot$	1.60 $R_\odot$	1.46 $R_\odot$
Luminosity	0.91 $L_\odot$	0.70 $L\odot$	1.43 $L_\odot$	1.01 $L_\odot$	2.76 $L_\odot$	2.01 $L_\odot$
Mass	1.00 $M_\odot$	1.00 $M_\odot$	1.00 $M_\odot$	1.00 $M_\odot$	1.00 $M_\odot$	1.00 $M_\odot$
Surface temperature	5888 K	5623 K	6056 K	5773 K	5877 K	5689 K

Stellar evolution

Stellar evolution in BiSEPS is determined by the use of fitting formulae, rather than a full evolution code such as the Dartmouth Stellar Evolution model (Dotter et al., 2008), Modules for Experiments in Stellar Astrophysics (MESA) (Paxton et al. (2011), Paxton et al. (2013), Paxton et al. (2015)) or the evolutionary tracks used in TRILEGAL (Girardi et al., 2000).

A full stellar evolution code provides much more information on internal processes within a star and their effect on parameters such as mass, radius, luminosity and temperature. With a sufficiently fine time resolution, details of the proportion of elements within the core, the temperature gradient and/or density gradient, for example, can be examined to understand the effect of small changes on the body as a whole. The fitting formulae, however, provide only a "snapshot" at a given evolutionary stage.

Incorporating models from a full evolutionary code, in particular MESA, into BiSEPS is a long term aim of those working with the code. In the long run, this will aid in the simulation of systems which have merged in the common envelope phase, a current limitation with BiSEPS, or in understanding the effect of tidally disrupted planet remains falling into a star, and its effect on stellar rotation, as discussed in Section 1.3 under asteroseismology.

However, for this project the fitting formulae are sufficient, as the time resolution of an evolutionary sequence captures changes in mass, radius, luminosity and temperature in sufficient detail. Given that observations of these parameters come with large error bars, something that PLATO 2.0 will help to address through its asteroseismological observations (Section 1.3), and that large areas of the Galaxy are being simulated, a "broad brush" approach is the only one that is computationally possible at this stage.

Model Galaxy

The model Galaxy used in BiSEPS is the simplest possible: a thin and thick disc, each described by a double exponential. The halo is not simulated, nor is the bulge. TRILEGAL offers a wider choice of disc models along with options for simulating the bulge and halo (Section 1.6.1), The Besançon model and its derivative, GUMS, have been carefully matched to observations and offer a sophisticated analysis of the thin disc, even incorporating stellar arms, as well as the thick disc, halo and bulge (Section 1.6.2). However, as will be shown in Section 2.3, the level of detail in BiSEPS is sufficient for good agreement between BiSEPS and GUMS.

As presented in Chapter 6, planets detectable by PLATO in a one-hour integration are restricted to host stars within 2 kpc (Fig. 6.16, Fig. 6.17, Fig. 6.18), as are unblended binaries (Chapter 7, Fig. 7.24), while contaminating blended binaries are restricted to $\approx$ 20 kpc from Earth (Chapter 7, Fig. 7.24).

The Galactic centre is at $l = 0°$. Fig. 1.9, from WISE mapping, indicates that the bulge is only observable between $l = 330°$ and $l = 30°$. Therefore the bulge not relevant to this work, as by line of sight it falls in neither LLN or LLS.

The halo will not contain exoplanet hosts or unblended eclipsing binaries observable by PLATO, due to the distances involved. Contaminating eclipsing binaries in the halo may be just about observable. The halo is sparsely populated, and this population is not evenly distributed: rather, it is concentrated in globular clusters. Future work verifying the accuracy of these simulations should take the halo, and its clumpiness, into account.

The results presented in Chapters 6-8 depend on an appropriate representation of the true stellar density and inhomogeneity. A further validation of the model stellar population using Gaia DR2 is an important future step.

Triples and higher hierarchies

At present, `BiSEPS` only models single stars and binaries, not triples and other hierarchies. Addressing this fact is recommended for future work (Chapter 9). In the course of this project, some preliminary work was done in understanding the issues involved.

Including triples and higher hierarchies in BiSEPS, a significant proportion of the observed population, will greatly enhance `BiSEPS` simulations. Indeed, triples represent $\approx 10\%$ of low mass stars and $\approx 50\%$ of B-class stars (Toonen et al., 2016). In effect, the calibration of the short period binaries described in Section 7.2.2 compensates for the fact that triples and higher hierarchies are not formally included. Raghavan et al. (2010) found that inner pairs in triple systems are more likely to be equal mass than the general population. The calibration of my simulations to observations of binaries with $P < 10$ days described in Section 7.2.2 finds that for this sample, binaries tend towards higher mass ratios. From Raghavan et al. (2010) it would therefore be reasonable to suggest that inner pairs of triples are a significant proportion of the short period binary population. The issue of triples and higher hierarchies should be addressed in future work, as my work-around cannot reflect the full picture. However, with $\approx$ 10% incidence, the overall impact of triples is likely to be limited. Stellar multiplicity also has a bearing on the evolution of systems, so needs to be considered carefully as part of the total population, and is computationally expensive. In addition, the longer periods in outer pairs in triples and higher hierarchies mean such systems are unlikely to have three eclipses during the timescale of a PLATO observation, so are unlikely to affect the results described in this work.

2.2.2 Advantages

`BiSEPS` offers a number of advantages over other population synthesis models for this project.

Perhaps the most significant is that the code is run locally rather than through an online interface. The user has complete control over the nature and format of the output. An example is the calibration by Farmer et al. (2013) for the synthetic *Kepler* field by adjusting the SFR to produce 1.2 star yr^{-1} with > 0.8 $M_{\odot}$, from the previous value of 1.0 star yr^{-1} with > 0.8 $M_{\odot}$.

Codes such as `TRILEGAL` and `Besançon` are very much centred on the perspective of an astronomer interpreting what is observed in the light of the best scientific knowledge of the Galactic composition. `BiSEPS` replicates this to some extent by only including systems observable from Earth at a given magnitude limit. However, with `BiSEPS` you know unambiguously which system is a binary and what the parameters of the two components are. Other codes may not disentangle observations of temperature and luminosity of the two components in a binary system, replicating the experience of an astronomer who sees a binary as a point source, and may interpret the binary as a single star. While the observer-oriented approach has many advantages in interpreting observations, for this particular thesis understanding the true nature of the binary population is much more significant. Therefore, `BiSEPS` is better suited than other codes to the work described here.

`BiSEPS` is straightforward to use. Its output matches observations and other population synthesis codes, as described in the following section, making it a reliable tool.

2.3 Calibration check

The `BiSEPS` fields generated for this work have been tested in terms of magnitude distribution and absolute numbers against data from Gaia DR1 (Gaia Collaboration et al. (2016), Lindegren et al. (2016), Arenou et al. (2017)), `GUMS` and the *Kepler* Input Catalogue from the Barbara A. Mikulski Archive for Space Telescopes (MAST) (Brown et al., 2011b) across two fields.

The first of these two fields covers a strip l = 64.5°-65.5°, b = 7.5°-54.5°, where the comparison is with `GUMS` and the Gaia DR1 output. The populations for `GUMS` and Gaia DR1 are reported in Gaia magnitudes G, and the population for BiSEPS is reported in Sloan magnitudes and converted from g', $r'-i'$ and $g'-r'$ to G using the two-colour formula from Section 5 of Jordi et al. (2010).

The second of the two fields is a synthetic sub-field drawn from the PLATO LLN simulation, approximating the on sky area observed by *Kepler* during its main mission. Here the comparison is with the *Kepler* Input Catalogue and Gaia DR1. The magnitudes of the *Kepler* Input Catalogue are all reported in Kp, and 82% are also reported in g', r' and i', enabling these to be converted to G using the two-colour formula from Jordi et al. (2010). For the rest of the *Kepler* population, an estimate is made of the appropriate proportion in each bin based on the outcome for the calculated sample. For example, there are 23,305 systems listed on MAST as having a *Kepler* magnitude in the range $12 \leq Kp < 13$. Of these, 22,114, or

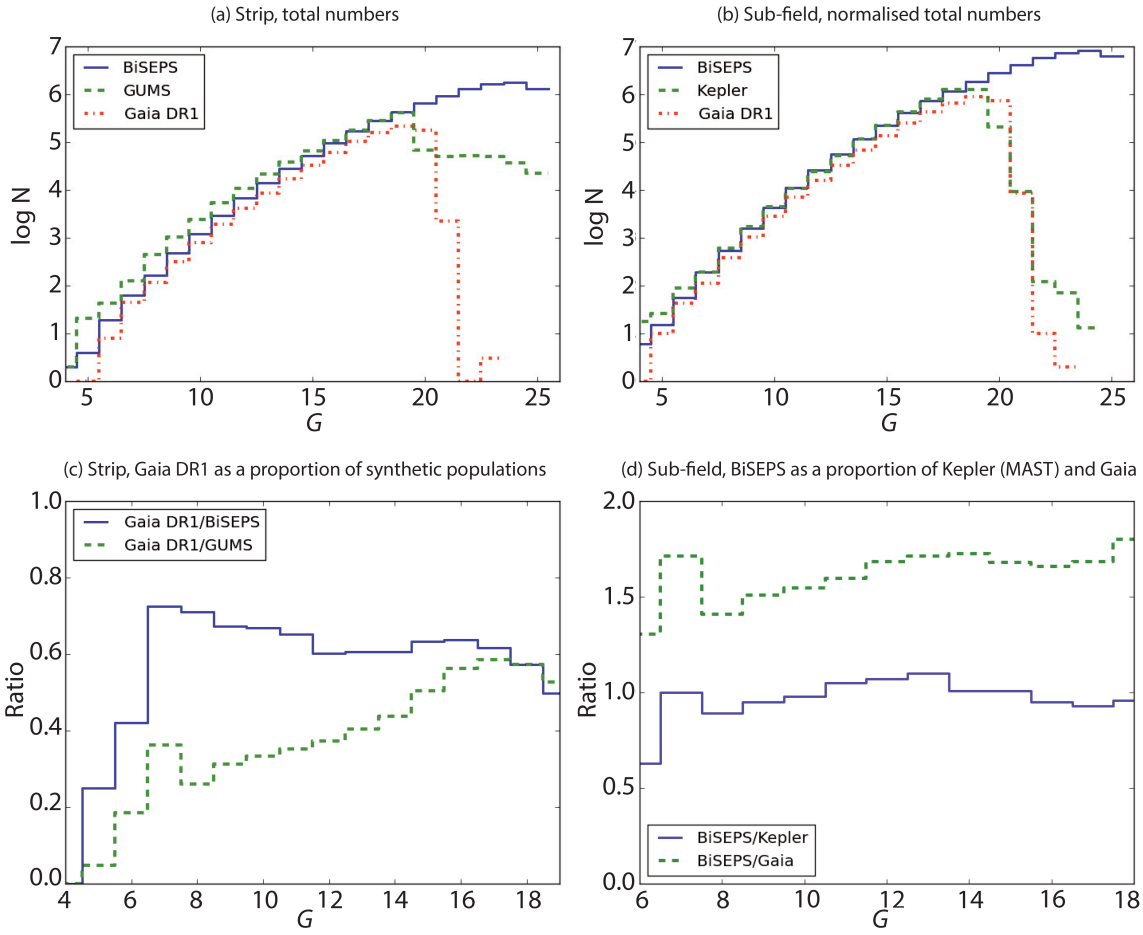


Fig. 2.3 Comparison of magnitude distribution in synthetic and observed populations. The sudden drop-off at faint magnitudes in observed populations in (a) and (b) is due to the observational limits of the surveys in question. (a) Binned total numbers in a strip l = 64.5°-65.5°, b = 7.5°-54.5°, `BiSEPS`, `GUMS` and Gaia DR1. (b) Normalised and binned total numbers in the subfield described in the text, `BiSEPS`, *Kepler* Input Catalogue and Gaia DR1. (c) In the strip described in (a), Gaia DR1:`BiSEPS` and Gaia DR1:`GUMS`. (d) In the subfield, normalised `BiSEPS` synthetic population:*Kepler* Input Catalogue (MAST) and BiSEPS synthetic population:Gaia DR1.

Table 2.3 Magnitude distribution converting from g', $r'-i'$ and $g'-r'$ to G using the two-colour formula from Section 5 of Jordi et al. (2010), for stars with $12 \leq Kp < 13$ on MAST.

Bin, G	Number	Percentage
< 4.0	3	0.014
4.0 to 5.0	1	0.005
5.0 to 6.0	1	0.005
6.0 to 7.0	0	0.000
7.0 to 8.0	1	0.005
8.0 to 9.0	1	0.005
9.0 to 10.0	1	0.005
10.0 to 11.0	13	0.059
11.0 to 12.0	739	3.342
12.0 to 13.0	21,350	96.545
13.0 to 14.0	4	0.018

94.9%, also have magnitudes recorded in all three of the g', r' and i' bands. The distribution of these 22,114 systems in G, when this is calculated using the two-colour formula, is shown in Table 2.3. The remaining 1,191 in the *Kepler* magnitude bin $12 \leq Kp < 13$ with only one, two or no Sloan magnitudes are then allocated to magnitude bins in the proportions suggested by those systems with all three Sloan magnitudes.

As the *Kepler* main mission observed 105 deg^2 (Prša et al., 2011) and the sub-field covers 194 deg^2, the results from `BiSEPS` and Gaia DR1 have been normalised by dividing the totals for `BiSEPS` and Gaia DR1 by $105/194 = 54.12\%$. The results for both fields are illustrated in Fig. 2.3.

In Fig. 2.3 (a) `BiSEPS` is closer to the Gaia DR1 population at brighter magnitudes and closer to the `GUMS` simulation at fainter magnitudes to about G = 20-21, beyond which the `GUMS` simulation becomes less complete. The Gaia DR1 observed population is consistently less than that predicted by both `GUMS` and `BiSEPS` (Fig. 2.3 (c)), even though `GUMS` was specifically created to simulate the expected Gaia observations.

In Fig. 2.3 (b), ignoring the brightest bins, which are affected by small number statistics, and the fainter bins, where coverage in both Gaia DR1 and the *Kepler* Input Catalogue are affected by detection limits, the `BiSEPS` synthetic population is in line with that expected from the *Kepler* Input Catalogue, while it would appear that Gaia DR1 lists $\approx 60\%$ of the expected number of sources, given that both the *Kepler* Input Catalogue and Gaia DR1 are based on observations. If both catalogues were similarly complete, it would be expected that the numbers of systems would be similar.

Gaia Collaboration et al. (2016) gives a number of reasons why Gaia DR1 is not complete. Most significant may be the effect of the scanning law: some parts of the sky will have been

visited much less frequently than others. Other issues described in Gaia Collaboration et al. (2016) affect bright stars, high proper motion stars, and extremely blue and red sources. Gaia Collaboration et al. (2016) stress in their Section 6 that DR1 is preliminary, based on "an incomplete reduction of a limited amount of raw Gaia data" and that all issues affecting completeness are expected to be addressed in future data releases.

Gaia DR2 was released in April 2018.[2] While it has not been possible to carry out a full calibration due to time constraints, some initial work has been done, taking advantage of the fact that, where a star in the *Kepler* Input Catalogue (KIC) is included in VizieR, Gaia DR2 can identified a star from its KIC number. A comparison between the KIC and Gaia DR2 is useful as both are based on observations.

Sixty Kepler Objects of Interest (KOIs), representing 52 stars, with estimates of stellar radii in both Gaia DR2 and the KIC were identified, and the ratio of the radii compared (Fig. 2.4). Beyond 1 kpc, apart from a rather obvious outlier, there is a clear trend that the further a star is from Earth (ie the smaller the parallax), the greater the discrepancy between the radii recorded in the two catalogues. Radii in the KIC are generally estimated from stellar tracks from the Dartmouth model, while radii in Gaia DR2 are derived purely from Gaia data (Andrae et al., 2018). The method described by Andrae et al. (2018) utilises the standard equation Eq. 1.2 to find the radius, which depends on T_{eff} and L. Andrae et al. (2018) described how T_{eff} is derived from the Gaia spectra, and for the stars in this sample there appears to be good agreement between the values recorded in the KIC and in Gaia DR2. L depends on a bolometric correction, derived from MARCS synthetic stellar spectra, and on the apparent magnitude, derived using the standard equation Eq. 1.4. While the apparent magnitude and parallax are observables, at present the authors state they are not making use of the extinction A derived from Gaia data, as *individual* extinction estimates are currently poor (their italics), so extinction is set at zero. This may explain the discrepancy between radii in Gaia DR2 and the KIC at > 1 kpc, and it is acknowledged in Andrae et al. (2018) that the radius and luminosity estimates are formally inconsistent with their extinction and reddening estimates.

Andrae et al. (2018) acknowledge that simply taking the inverse of the parallax can give biaised results in Gaia data, but respond to this by only calculating luminosities, and hence radii, where the fractional uncertainty in the parallax is < 0.2: so while 161 million stars have estimates of T_{eff}, only 77 million have estimates of L. This condition would therefore apply to all the stars I have used in my test sample. Binaries are also excluded from Gaia DR2, although some close binaries will be recorded as point sources, and some wide binaries may have parameters of both stars included as single resolvable sources. Full coverage of the binary population is expected in Gaia DR3.

[2]https://gea.esac.esa.int/archive/

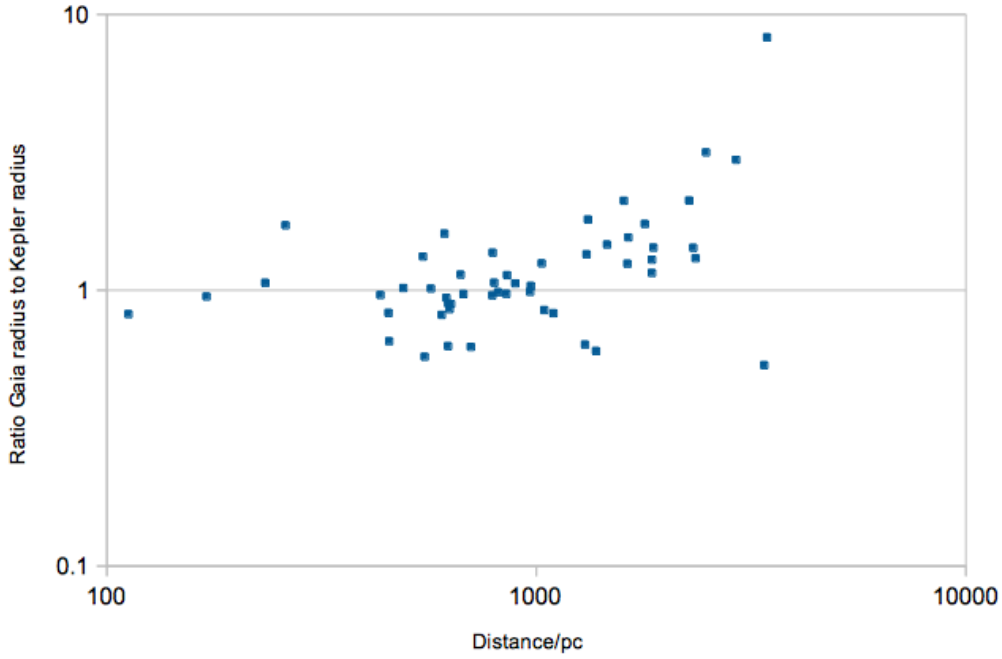


Fig. 2.4 Ratio of the radii of a selected number of stars as recorded in Gaia DR2 to the value recorded in the *Kepler* Input Catalogue as a function of distance, as derived from the parallax recorded in Gaia DR2.

Comparison of BiSEPS with Gaia DR2 will be illuminating in terms of magnitude of all systems, distance to and effective temperature of all single stars, and physical parameters (where recorded) of stars < 1 kpc. However, comparison of luminosities and radii of the stellar population at a distance > 1 kpc should wait until there is increased confidence by the Gaia collaboration in the individual extinctions of stars.

2.4 Summary: Chapter 2

In this chapter, I have discussed the history of BiSEPS, its advantages and its limitations. The stellar population models, based on observational data, have been well tested over time in a variety of applications, by those who used the code before me. Nevertheless, I have conducted my own calibration checks both against other synthetic populations, specifically GUMS, and observations recorded in the *Kepler* Input Catalogue (KIC) and first Gaia data release. Early comparisons with Gaia DR2 have been made, although further work is required for a full comparison.

Having generated a synthetic binary population, an essential step in understanding the eclipsing binary part of this population is the generation of synthetic light curves. The following chapter describes how this is achieved in this project.

Chapter 3

Synthetic light curves

The estimated primary and secondary eclipse depths in an eclipsing binary system are derived in this project using a version of JKTEBOP (Southworth et al. (2004), Southworth et al. (2005), Southworth et al. (2007), Southworth (2008), Southworth et al. (2009), Southworth (2010), Southworth (2011), Southworth (2013)), integrated into BiSEPS by Robert Farmer, developed by Enda Farrell and further refined by myself.

3.1 Principles of synthetic light curve generation

As described on the site through which the code is distributed[1], JKTEBOP calculates light curves by numerical integration of concentric circles over each star, modeling stars as spheres when calculating eclipse shapes and as biaxial ellipsoids when calculating proximity effects.

The input required is absolute magnitudes, radii, mass, limb darkening coefficients and gravity darkening coefficients of the two stars, along with eccentricity, semi-major axis, angle of inclination and orbital period of the system.

Each transit is assessed at 6,000 points, giving a resolution of $\Delta\phi = 1.67\ 10^{-4}$, or 0.06° (3.6 arcmin) measured from the centre of mass of the binary system.

3.1.1 Limb darkening

Although photons may be emitted from a star at any angle, those most likely to reach an observer are emitted from the centre of an observed stellar disc rather than from the edge, or limb (Fig. 3.1). If an imaginary sphere with unit radius is placed around the point on the stellar disc at which the photon is emitted, the subset of orientations at which this is emitted can be seen as drawing a small circle on the surface of the sphere.

[1] http://www.astro.keele.ac.uk/jkt/codes/jktebop.html

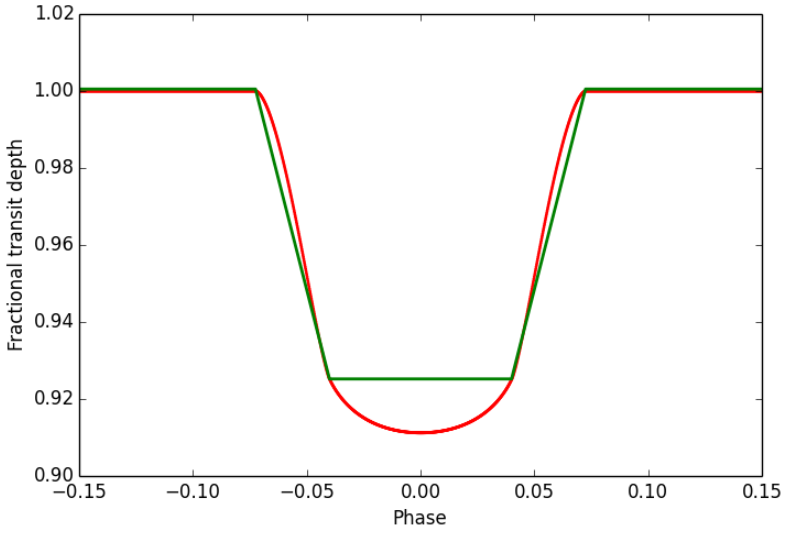


Fig. 3.1 Illustrating the effect of limb darkening. Green: an eclipse without taking account of limb darkening. Red: an eclipse taking limb darkening into account.

When $\theta = 0°$, with $\theta = 0°$ set as the vector from the star to the observer and not the vector perpendicular to the surface of the star, this small circle becomes a point and the subset of orientations is at a minimum. When $\theta = 90°$, it becomes a great circle and the subset of orientations is at a maximum: it is also unobservable by the observer to whom the vector at $\theta = 0°$ is pointing.

In spherical co-ordinates, the area of the small circle with unit radius and width $\Delta\theta$ is:

$$\int_0^{2\pi} \sin\theta \Delta\theta \mathrm{d}\phi = 2\pi \sin\theta \Delta\theta \tag{3.1}$$

The probability of observing a photon emitted at a given angle θ is the cross product of the two vectors, and is therefore sinusoidal. Since the surface area of a hemisphere of unit radius is 2π, the probability that θ will be in the range $[\theta', \theta' + \mathrm{d}\theta]$ is

$$\frac{2\pi \sin\theta \Delta\theta}{2\pi} = \sin\theta \Delta\theta \tag{3.2}$$

The probability that a given angle θ' at which a photon is emitted will be found in the range $\Delta\theta'$ is

$$P = \int_{\theta}^{\prime\theta'+\Delta\theta'} \sin\theta \mathrm{d}\theta = [-\cos\theta]_{\theta}^{\prime\theta'+\Delta\theta'} \tag{3.3}$$

The angle between the line of sight to the observer and the direction of travel is γ. At the observed centre of the stellar disc, the probability that γ is sufficiently low to be observed is signficantly higher than on the limb of the stellar disc.

This phenomenon is known as limb darkening, and various models are available to describe this, including linear, quadratic, square root, logarithmic, cubic and Claret's four-parameter law (Wade and Rucinski (1985), van Hamme (1993), Claret (1998), Claret (2000), Barban et al. (2003), Claret and Bloemen (2011)). JKTEBOP can work with any of the above. In this work, the quadratic limb darkening law is used. Where I is intensity, and u and v are co-efficients:

$$\frac{I(\cos(\gamma))}{I(\cos(0))} = 1 - u(1-\cos(\gamma)) - v(1-cos(\gamma))^2 \tag{3.4}$$

Because JKTEBOP integrates concentric circles, centred on the observer's line of sight, γ can be reset as a constant for each concentric circle, ensuring accurate analysis of limb darkening effects, given a sufficiently fine resolution.

3.1.2 Gravity darkening

Gravity darkening occurs when a star rotates so rapidly that its shape is oblate rather than spherical, due to centrifugal forces (Fig. 3.2). Because stars spin most rapidly at birth, and because stellar mass is also a factor, gravity darkening is more significant in young, massive stars. In such stars, the equator appears darker than expected relative to the poles and the effect will be most obvious on the limb of the equator.

Gravity darkening has been modelled in follow up observations of *Kepler* objects of interest. In a study on KIC 5006817, Beck et al. (2014) used asteroseismology to study a pulsating red giant star in an eccentric binary system. Beck et al. (2014) state that the rotation period of the envelope is at least 165 days, roughly twice the orbital period, the stellar core rotates 13 times faster than the surface, and that the gravity darkening exponent is "larger than expected". Barnes et al. (2011) measured the spin-orbit misalignment of KOI13.01 from its gravity-darkened *Kepler* light curve, and found that the transit fits with and without gravity darkening are very similar. Barnes et al. (2011) therefore concluded that the effect of gravity darkening on a transit light curve is negligible.

Gravity darkening is not expected to be significant in the majority of the binary systems included in this study, ie MS/MS systems, but is important in those rarer binaries incorporating at least one massive, young star and those in close proximity.

JKTEBOP uses both gravity darkening and limb darkening co-efficients to calculate the minimum and maximum flux from each star, F_{min} and F_{max}. When the star is not oblate, $F_{min} = F_{max}$. F_{min} and F_{max} are used to find a coefficient, δ

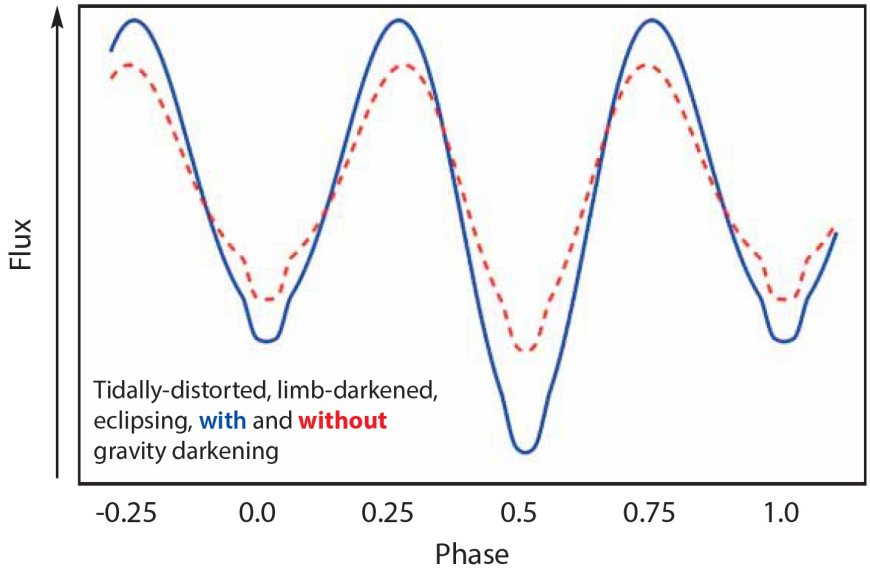


Fig. 3.2 The effect that the presence or absence of gravity darkening will have on the light curve of an eclipsing binary. Red: no gravity darkening. Blue: with gravity darkening.

$$\delta = 1 - \frac{F_{min}}{F_{max}} \tag{3.5}$$

which is zero where the star is not oblate, and which in turn is used to find the brightness contribution ($C_i (i = 1, 2)$) from each component:

$$C_1 = \frac{\pi r_{b,1}^2 F_{max,1}}{B_1 + B_2}(1 - \delta_1 S) \tag{3.6}$$

$$C_2 = \frac{\pi r_{b,2}^2 F_{max,2} S_B}{B_1 + B_2}(1 - \delta_2 S) \tag{3.7}$$

where r_b is the semi-minor axis of the star, B_i is the un-normalised brightness of each stellar component, S_B is the surface brightness ratio and S is a coefficient which depends on phase ϕ, the tidal lead angle τ and the angle of inclination i:

$$S = sin(i)^2 (cos(\tau)cos(\phi) - sin(\tau)sin(\phi))^2 \tag{3.8}$$

Where $\tau = 0$, this simplifies to

$$S = sin(i)^2 cos(\phi)^2 \tag{3.9}$$

So JKTEBOP incorporates oblateness arising from gravity darkening, as well as distortion of the shape of the star due to a tidal lead angle when the star is oblate, automatically into its

calculations of the brightness contribution from each star. In this work, most binaries are not affected by gravity darkening and are spherical rather than oblate.

3.1.3 Area eclipsed

As shown in Eq 1.10 (Section 1.5), the in transit flux is defined by the flux of the eclipsing star plus the flux of the uneclipsed area of the star being eclipsed, over the total flux of both stars. If the eclipse is grazing, the area that is eclipsed is not a circle.

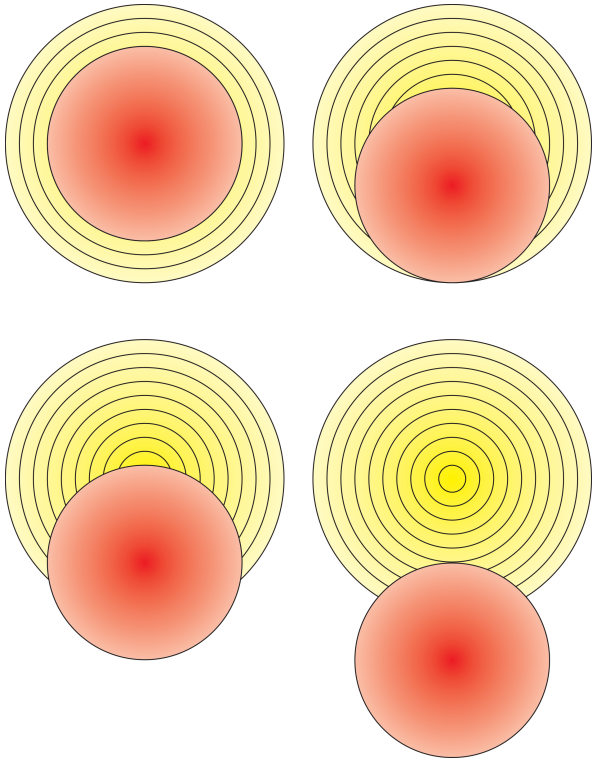


Fig. 3.3 Cartoon illustrating the method of calculating limb darkening effects and area eclipsed using concentric circles, at $\phi = 0$. The angle of inclination depends on the radii of the two stars and the semi-major axis, a. Assuming R1 = 1 $R_\odot$, R2 = 0.7 $R_\odot$ and a = 100 $R_\odot$, the inclinations are: top left 90°, top right 89.8°, bottom left 89.6°, bottom right 89.3°, with a critical angle of 89.0°. Assuming a = 10 $R_\odot$, the angles of inclination are top left 90°, top right 88.3°, bottom left 86.0°, bottom right 82.5° and the critical cangle is 80.2°.

The concentric circle approach is primarily used for its advantages in limb darkening, as described above, but is also useful in calculating the area eclipsed. Fig 3.3 illustrates this. Using a mesh that is far finer than that illustrated in the cartoon, as well as phase ϕ, the angle of inclination i and the radii of the two stars R_1 and R_2, the proportion of each concentric circle that is eclipsed can be calculated, giving, when integrated, A_2, the eclipsed area. A_1 is simply πR_1^2 where the star is spherical, as is the assumption in JKTEBOP.

3.1.4 Ellipsoidal variations

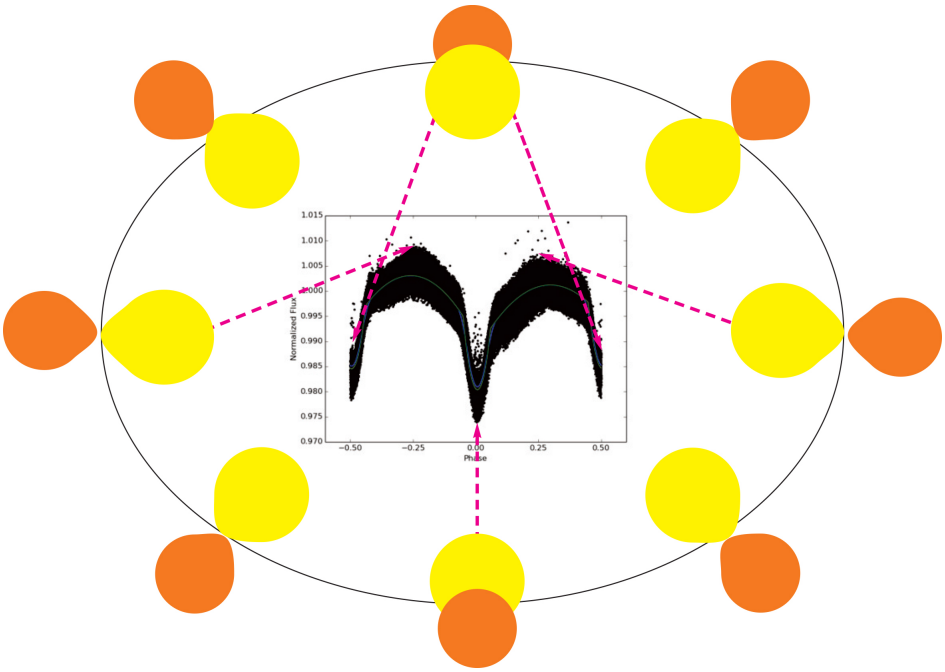


Fig. 3.4 Cartoon illustrating the effect of ellipsoidal variations on a light curve. Example light curve is KIC 6353203 from the *Kepler* Eclipsing Binary Catalogue. Colours and sizes are representative to emphasise the relationship between ellipsoidal variations and phase, and are not intended to accurately represent the stars in KIC 6353203.

Fig 1.7 demonstrated, from light curves of eclipsing binaries with $P < 1$ day from the *Kepler* Eclipsing Binary Catalogue, how the presence or absence of ellipsoidal variations can make a shallow eclipsing binary appear less or more planet-like. To discuss the importance of ellipsoidal variations, I take the example KIC 6353203 from Fig 1.7, which includes evidence of both a clear transit and ellipsoidal variations, and demonstrate the orbital configurations which give rise to this particular light curve in Fig 3.4. I stress that Fig 3.4 is a cartoon, and that the colours and relative sizes have been chosen to emphasise the point, not to be realistic.

It is likely that one star is more luminous than the other, as the eclipse at $\phi = 0.5$ is shallower than that at $\phi = 0$. The relative eclipse depth is ≈ 0.025 for the primary eclipse and ≈ 0.020 for the secondary eclipse, including both the transit and the ellipsoidal variations, and is measured relative to $\phi = n/8$ where n is odd. The maxima at $\phi = 0.25$ and $\phi = 0.75$ are ≈ 0.005 *above* the reference point for the normalised flux at one-eighth phase. The light curve appears noisy: this is a result of the shallowness of the eclipse. The eclipse itself is sharply defined.

The short orbital period of 0.509 days indicates that the two stars are close to one another. One may be close to filling its Roche lobe. Tidal effects are distorting the shape of both stars,

pulling the surfaces of both stars closer to one another in a tidal bulge. At $\phi = \pm 0.25$, when the stars are in profile from the perspective of the observer, they are at their most distorted and at maximum brightness: at $\phi = 0$ and $\phi = 0.5$ the tidal bulges are most likely to be hidden from the observer and the stars are most likely to appear spherical during an eclipse. Even if the stars did not eclipse one another, they would be at minimum brightness at $\phi = 0$ and $\phi = 0.5$: the two examples in Fig 1.7 with the highest morphology parameter (KIC 10417986 and KIC 6224853), illustrate this as the light curves are sinusoidal with equal depth minima, while KIC 5128972, which also has a high morphology parameter, is suggestive of an eclipsing binary only in the unequal minima, as a transit-like dip in the light curve is not visible by eye.

The profiles offered by KIC 6353203 at ϕ = 0.125, 0.375, 0.625 and $\phi = 0.875$ are similar in all cases.

JKTEBOP calculates ellipsoidal variations by treating each star as a biaxial ellipsoid, adjusting the exact shape of the ellipsoid according to the precise phase. Where stars in a binary are sufficiently separated for tidal bulges to be negligible, ellipsoidal variations will be insignificant. "Sufficiently separated" does, of course, depend on the evolutionary state of the binary: a red giant, for example, expands significantly, making a tidal bulge and eventual mass transfer more likely than when the same star was on the main sequence.

For a summary of the gravitational processes that occur as a star fills its Roche lobe, see the subsections on *Semi-detached* and *Contact or overcontact* systems in Chapter 1.5.

JKTEBOP enables ellipsoidal variations to be accounted for and quantified in a synthetic eclipsing binary population, and also to be disregarded when required.

3.2 Application in BiSEPS

JKTEBOP is used to create look-up tables which can be consulted for a known BiSEPS model at a given angle of inclination. These tables, saved in an SQL database, include BiSEPS reference numbers, primary and secondary depths with and without ellipsoidal variations, angle of inclination, critical angle, orbital period and noise. Two BiSEPS reference numbers are included, model number and line number and these two, with metallicity, enable a system in the synthetic galaxy to be uniquely matched to the relevant models in the SQL database, thereby saving disc space by not requiring other parameters such as mass, radius, temperature and luminosity of the two stars to be saved in the SQL database. The primary eclipse depth is assumed to be that in which the evolutionary primary is occluded although, in systems in which one star is a compact remnant, this may in fact be the shallower depth: this is taken into account when results are interpreted. The development of these tables was carried out by Enda Farrell, building on work started by Rob Farmer.

I was working with a series of fields each no more than 1° by 1°. Each field required its own JKTEBOP table, with one set of eclipse depths (at a resolution of inclination $i = 0.1°$) for each star in the field. As a result, some of the SQL databases could become quite large, especially closer to the Galactic plane. However, the alternative approach of generating master SQL tables from the master list of BiSEPS models from which the synthetic populations were drawn, would have led to greater problems. The maximum SQL database which could be loaded into memory when running the simulation code is close to 1 GB. The master tables would have far exceeded that size, so would need to be divided into manageable chunks. The most time consuming part of running the simulation code, in proportion to results achieved, was the loading of the SQL database. Therefore, being able to load it once for each 1° by 1° field was the most efficient way of working. Also, through testing, it was found that each SQL database needed to be generated from the field in question: while a field generated from a dense field might include multiple copies of the common model systems, rare model systems may be missing.

For a given binary in a synthetic population, a preliminary code first generates those prerequisites described in section 3.1 which are not already available: namely the absolute magnitudes of both stars in the *Kepler* passband, the quadratic limb darkening coefficients and the gravity darkening coefficients of the two stars. The absolute magnitudes are derived from the apparent magnitudes, luminosities, effective temperatures and surface gravities of each star in the system, compared with reference values for the Sun, utilising standard equations. Extinction is also recalculated. Gravity darkening coefficients and quadratic limb darkening coefficients are calculated from the effective temperature, the surface gravity and the metallicity of each star, taking into account, as with the absolute magnitudes, the evolutionary class of the star.

The critical angle, i_c, is calculated: this is the minimum angle at which an eclipse occurs and, where a is the separation between the stars, is

$$i_c = \arccos\left(\frac{R_1 + R_2}{a}\right) \tag{3.10}$$

At an angle of inclination below the critical angle no eclipse can be observed. The version of JKTEBOP which I inherited is set to ignore angles of inclination below the critical angle for a given system. This means that non-transiting ellipsoidal binaries with highly sinusoidal light curves analogous to systems in the *Kepler* Eclipsing Binary Catalogue with a morphology parameter > 0.8 (see Section 1.5) which are clearly not transiting exoplanets are not included in our simulations.

JKTEBOP then generates a series of models from the critical angle to 90° (edge on), deriving the relative eclipse depth at $\phi = 0$, $\phi = 0.25$ and first contact, the point at which

the secondary stellar disc first "touches" the primary stellar disc, from the perspective of the observer. Relevant information is then saved to a look-up table, as described in the first paragraph in this subsection.

From Haswell (2010), when eccentricity $e = 0$ the transit duration is

$$T_{dur} = \frac{P}{\pi} \arcsin\left(\frac{\sqrt{(R_1+R_2)^2 - a^2\cos(i)^2}}{a}\right) \tag{3.11}$$

Normalising the orbital period, and remembering that to find the point of first contact and of fourth contact ϕ_{FC}, which are symmetrically distributed around $\phi = 0$, we only need consider half the orbital period and hence half the transit duration, leads to

$$\phi_{FC} = \frac{1}{2\pi} \arcsin\left(\frac{\sqrt{(R_1+R_2)^2 - a^2\cos(i)^2}}{a}\right) \tag{3.12}$$

This equation is valid in the case where, as with `BiSEPS` at present, eccentricity is zero. If, in future work as proposed in Chapter 9, eccentricity is considered within BiSEPS, then Eq 3.11 would become

$$T_{dur} = \frac{P\sqrt{1-e^2}}{\pi(\sin(\theta+\omega_{OP}) + e\sin\omega_{OP})} \arcsin\left(\frac{\sqrt{(R_1+R_2)^2 - a^2\cos(i)^2}}{a}\right) \tag{3.13}$$

and Eq 3.12 would become

$$\phi_{FC} = \frac{\sqrt{1-e^2}}{2\pi(\sin(\theta+\omega_{OP}) + e\sin\omega_{OP})} \arcsin\left(\frac{\sqrt{(R_1+R_2)^2 - a^2\cos(i)^2}}{a}\right) \tag{3.14}$$

where θ is the "true anomaly" which measures how far around the orbit from the pericentre the planet has travelled, the pericentre being the point at which the planet is closest to the star, and ω_{OP} measures the orientation of the pericentre with respect to γ, where γ is the intersection of the orbit with the positive x axis, where the x axis is defined by the intersection of the orbit with the plane of the sky as seen by the observer and positive x is on the side of the orbit where the planet moves towards the observer (Haswell, 2010). If the depth due to the eclipse is less significant than the sinusoidal depth due to ellipsoidal variations (see, for example, panel d in Fig 1.7, KIC 6353203), the system is unlikely to be mistaken for a transiting exoplanet.

3.3 Summary: Chapter 3

This work uses JKTEBOP to derive synthetic light curves and hence determine the fractional decrease in flux in eclipsing binaries, as adapted for use with `BiSEPS` by previous members of my group. The code is most suitable for use in detached and semi-detached binaries, and these will dominate the proportion of eclipsing binaries that may be mistaken for transiting exoplanets. The code takes into account limb darkening and, where appropriate, gravity darkening and ellipsoidal variations.

Eclipse depths are saved to SQL databases, which are utilised in the eclipsing binary simulations (Chapter 7).

Having described the codes used in this project, in the next chapter I will describe the synthetic Galaxy simulated for this work.

Chapter 4

Galaxy Simulation

Integral to this work is the synthetic Galaxy simulated using BiSEPS. This chapter describes the on-sky region that has been simulated, and outlines its use in the eclipsing binary and exoplanet simulations, particularly examining the binary fraction.

As indicated in Chapter 2, BiSEPS has previously been used to simulate on sky fields, such as the *Kepler* field simulated for and studied in Farmer et al. (2013). During that process, the parameters required to match observations, such as star formation rate, were carefully calibrated. The result was a code that could be applied to investigations of much larger fields, such as those PLATO will survey.

Table 4.1 lists the fields used in this work, the relevant chapters and the purpose for which the field was used.

4.1 On-Sky Regions Simulated in this Study

A considerable portion of the on sky regions that will be observable by PLATO was simulated during the course of this study. These regions are indicated in the green boxes in Fig. 4.1, overlain on a diagram from page 83 of the PLATO Definition Study Report, 4 April 2017[1] indicating the present thinking with regards to the positions of the two Long Look fields (Rauer et al., 2014) first described in Section 1.3 of this thesis, and the potential Step & Stare fields. The areas covering the Long Look fields were studied in detail. The remaining regions were held for future work on the Step & Stare fields, although it should be noted that the bulge has not been accounted for, should the synthetic Step & Stare fields require study of the region between $l = 330^\circ$ and $l = 30^\circ$. Moving closer to the Galactic plane will be computationally expensive, in terms both of disk space and time.

[1] http://sci.esa.int/jump.cfm?oid=59252

Table 4.1 Summary of synthetic fields utilised in this work

Description	Chapters	Purpose
Strip, 64.5° < l < 65.5°, 7.5° < b < 54.5	2.3	Calibration check
194° sq region including the Kepler field	2.3	Calibration check
	4.2	Calibration check
	6.3	Calibration check
Arc, 240° < l < 270°, -54.5° < l < -5.5°	4.2	Calibration check
Synthetic *Kepler* field from Farmer et al. (2013)	4.2	Calibration check
	6.1	Generate intrinsic exoplanet distribution
	7.2	Calibrate eclipsing binaries by mass ratio
	7.3	Calibrate eclipsing binaries by period
Approximation of Long Look North field	6.4	Synthetic exoplanet population
	7.4	Synthetic eclipsing binary population
Approximation of Long Look South field	6.4	Synthetic exoplanet population
	7.4	Synthetic eclipsing binary population
1 deg sq, centred on l = 50°, b = 8°	7.2	Calibrate eclipsing binaries by mass ratio

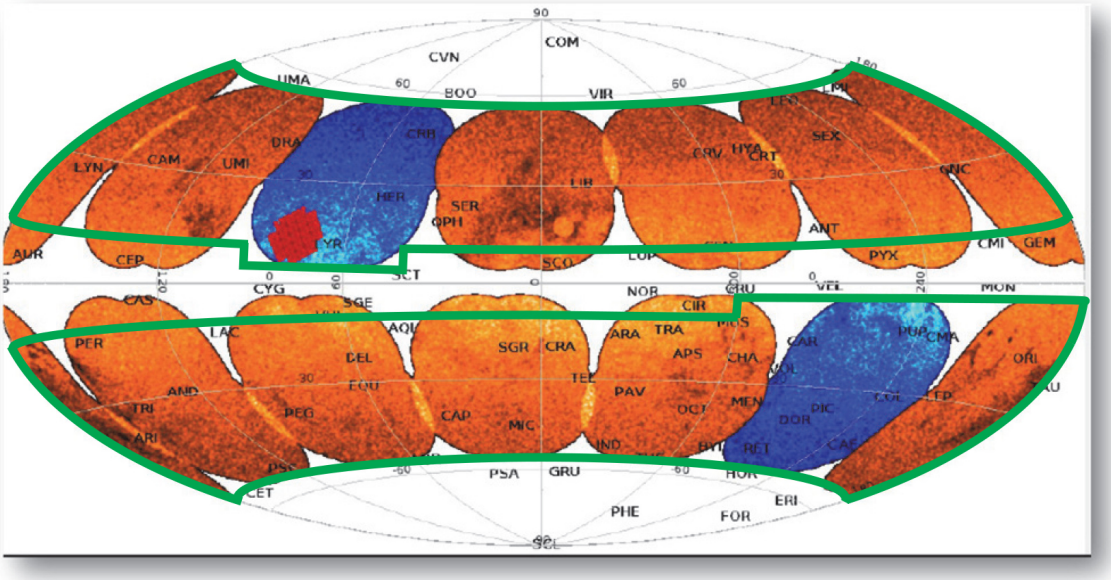


Fig. 4.1 Annotated map adapted from the PLATO Definition Study Report, 4 April 2017 (http://sci.esa.int/jump.cfm?oid=59252), page 83: all regions simulated in this study are enclosed in the green boxes, while the regions studied in greatest detail approximately match the blue Long Look fields. The region in red, the *Kepler* main mission, is the region studied in detail in Farmer et al. (2013). The orange fields represent alternatives for Step & Stare fields.

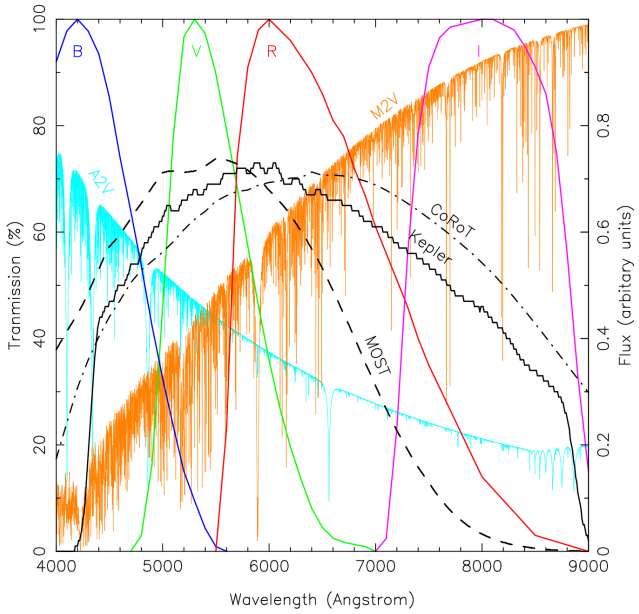


Fig. 4.2 Selected transmission functions and passbands, from Rowe et al. (2009). The transmission functions for the Johnson filters have been scaled to peak at 100% transmission. The spectrum for an A2V star (cyan) and for a M2V star (orange) have been scaled to have equal flux in the Johnson V filter. MOST (Microvariablity and Oscillations of STars) is a Canadian satellite observing bright transiting exoplanet systems.

At all values of l, two bands were simulated: $-54.5^\circ < b < -11.5^\circ$ and $8.5^\circ < b < 54.5^\circ$. Additionally, as shown in Fig. 4.1, regions covering the proposed Long Look fields were simulated closer to the Galactic plane, to $b = \pm 5.5$. The region $40.5^\circ < l < 89.5^\circ$, $5.5^\circ < b < 54.5^\circ$ was used to approximate the Long Look North (LLN) field and the region $228.5^\circ < l < 277.5^\circ$, $-54.5^\circ < b < -5.5^\circ$ was used to approximate the Long Look South (LLS) field. These regions are approximations, following lines of Galactic longitude and latitude: the final position and orientation of PLATO is work in progress. In addition, when determining which systems would be observable by PLATO, I was able to make use of code, provided by Catala in a private communication, determining the field of view of each group of cameras and the regions in which groups of cameras would overlap (Fig. 1.5).

The LLN field contains within it the field observed in the *Kepler* main mission. As described in Section 2.3, the *Kepler* Input Catalogue was utilised to verify the magnitude distribution of a synthetic on-sky population in a region encompassing that observed in the *Kepler* main mission.

The synthetic population has a magnitude limit of $Kp \leq 26$. The *Kepler* passband is $420 \leq \lambda \leq 890$ nm (Brown et al., 2011b), while the PLATO Definition Study Report states that the normal cameras on PLATO are expected to operate in $500 \leq \lambda \leq 1000$ nm.

The faintest stars included in the PLATO priority populations are in P4, M class stars, $V < 16$. This is the only class observing stars fainter than $V = 13$. Fig. 4.2, from Rowe et al. (2009), demonstrates that an M-class star, which peaks in the infrared, is fainter in the V band than it is in the redder part of the *Kepler* band, remembering of course that the flux of both the A2V star and the M2V star in this figure have been normalised to be equal at the peak of the V band.

The faintest background eclipsing binaries expected to survive blending and give an Earth-like signal will be no more than ≈ 10 magnitudes fainter than the star that is being observed, where "survive blending" means still be observable after blending is taken into account. This is based on the assumption that the deepest eclipse depths are 50%: as shown in Chapter 7, this assumption is not always true once limb darkening and mass transfer are taken into account, but this is what the standard simplification is based on so we will let it hold for now. That simplification is based on two equal mass, equal radius stars eclipsing one another at an angle of inclination of 90°. Referring to Eq 1.14 and assuming that the actual eclipse depth is 0.5, we find that when $\Delta m = 1$, the observed depth is ≈ 0.2 or 200,000 ppm, still clearly stellar. When $\Delta m = 5$, the observed depth is ≈ 0.005 or 5,000 ppm, possibly a giant planet. When $\Delta m = 10$, the observed depth is ≈ 0.00005, or 50 ppm, possibly a terrestrial planet. As will be discussed in Chapter 5, the target for the central part of the field, where all four groups of cameras are focused, is 800 ppm for the P4 class. This would be achieved in this example with $\Delta m \approx 9.5$. The target for the central part of the field for P1 is 34 ppm: however, the magnitude limit for P1 is $V = 11$, so background blends with P4 targets may be fainter than background blends with P1 targets.

These considerations led to the selection of the magnitude limit in this work. A brighter limit would have reduced the demands on disk space, and therefore would have allowed simulations in additional regions to be carried out closer to the Galactic plane. However, completeness in the region that was studied was considered more important.

4.2 Use in Eclipsing Binary and Exoplanet Simulations

Once the fields had been simulated, two sets of synthetic samples were generated: a synthetic eclipsing binary distribution, using light curves generated in JKTEBOP, and a synthetic exoplanet distribution.

Kepler observations were used to calibrate these two distributions: this is described in more detail in Chapters 7 and 6 respectively, but in essence the publicly available data on unblended eclipsing binaries (*Kepler* Eclipsing Binary Catalogue) and on confirmed exoplanets (NExSci) was used to calibrate the synthetic samples obtained from the synthetic field described in Farmer et al. (2013), and then apply these calibrations to synthetic samples obtained from the

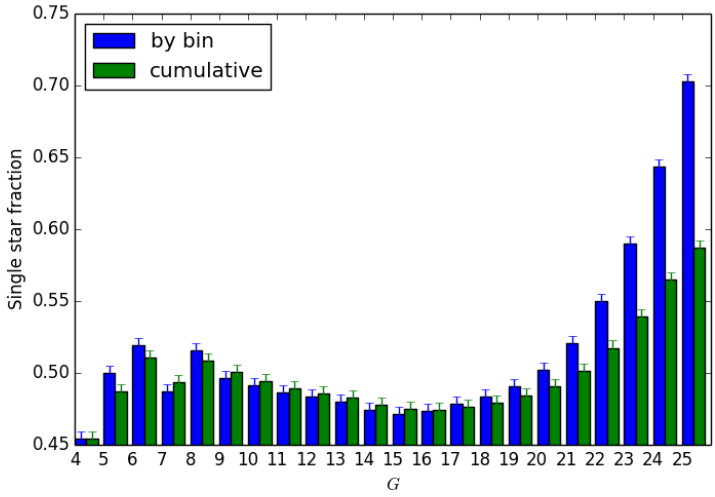


Fig. 4.3 The fraction of single stars in a synthetic field with an initial binary fraction of 0.5, approximating the region surveyed in the *Kepler* main mission. Blue: single star fraction as a function of apparent magnitude G. Green: single star fraction in an ensemble of systems with magnitude $< G$ (the so-called cumulative single star fraction).

larger scale synthetic fields generated for this work. Once the synthetic eclipsing binary and exoplanet samples were obtained, the ratio of planets to eclipsing binaries could be calculated.

Analysing a synthetic field from the PLATO simulations which encompassed the equivalent area to that observed by the *Kepler* main mission, it was found that the observed binary fraction depends on apparent magnitude. Fig. 4.3 illustrates the fraction of single systems, as a function of G. Also shown is the single star fraction in the ensemble of systems with magnitude $< G$, referred to for conciseness in the figure legend as "cumulative". G is used as these figures were obtained as part of the calibration process described in Chapter 2.3, where the synthetic field was compared with observations from the *Kepler* Input Catalogue hosted by MAST, and Gaia DR1.

As it had been observed that squares of equivalent latitude in the Northern and Southern hemispheres, for example $b = \pm 19.5°$ to $b = \pm 20.5°$, contained approximately 11% more stars in the South than the North at equal values of l, a similar process was carried out for the central portion of the synthetic Long Look South field, for the region $-54.5° < b < -5.5°$, $240° < l < 270°$. Results were binned both by G and by b. The results in this field by b indicate that the ensemble of systems with magnitude $G < 26$ had a single star fraction of 0.54 at all values of b, while the ensemble of systems with magnitude $G < 16$ had a single star fraction of 0.51 closer to the Galactic South pole and 0.49 closer to the Galactic plane. The dependence on latitude was therefore weak. The results by G are illustrated in Fig. 4.4.

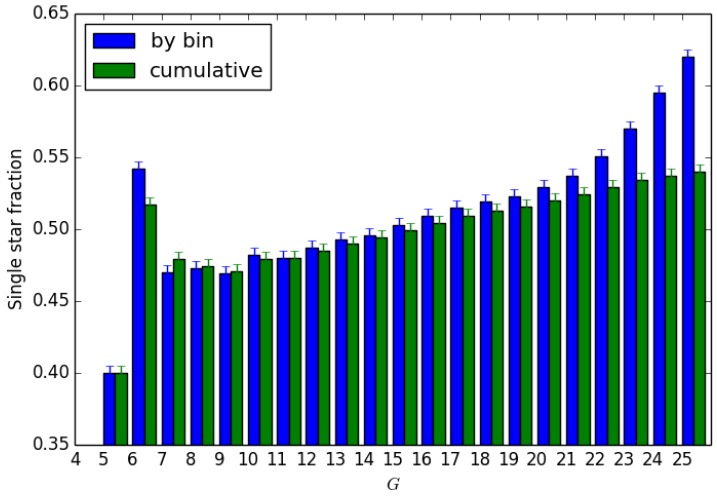


Fig. 4.4 Single star fraction in a synthetic population with an initial binary fraction of 0.5, in the central part of the synthetic Long Look South field, $-54.5^\circ < b < -5.5^\circ$, $240^\circ < l < 270^\circ$. Blue: single star fraction as a function of apparent magnitude G. Green: single star fraction in an ensemble of systems with magnitude $< G$ (the so-called cumulative single star fraction).

Disregarding the peak at the bright end of the spectrum, which is probably due to small number statistics, these results provide an interesting comparison with Fig. 4.3. In Fig. 4.4 there is a steady increase in the single star fraction in both sets of data from $G \approx 8$: in Fig. 4.3 there is a decrease in the single star fraction from $G \approx 8$ to $G \approx 15$, before a steady increase to fainter magnitudes: this increase being sharper than the one noted in Fig. 4.4. This indicates that there may be a real difference in the populations in the synthetic LLS and LLN fields used for this study.

Aside from the fact that the LLS field looks through the Galactic plane, where the LLN field does not, another contributory factor may be the initial binary fraction. In nature, this is dependent on stellar mass: in BiSEPS, it is not. In nature, as described in Raghavan et al. (2010), for example, the more massive the star, the more likely it is to have a stellar companion, while in BiSEPS the binary population is seeded regardless of stellar mass of the primary. Addressing this is one of the issues described under Future Work in Chapter 9. In the meantime, it can be expected to lead to an over representation of low mass binaries in the synthetic population.

In the calibration field in LLN, a total of 59% of the systems in the calibration field are single stars where $4 < G < 26$, 49% where $4 < G < 21$ and 47% where $4 < G < 16$.

The fields had been generated with a magnitude limit in Kp, while the G magnitude is found by conversion from the g', r' and i' magnitudes. To verify that the features noted were

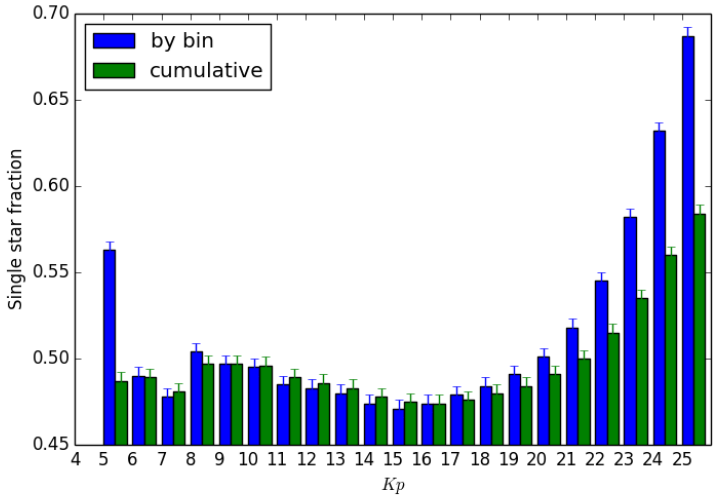


Fig. 4.5 The fraction of single stars in a synthetic field with an initial binary fraction of 0.5, approximating the region surveyed in the *Kepler* main mission. Blue: single star fraction as a function of apparent magnitude Kp. Green: single star fraction in an ensemble of systems with magnitude $< Kp$ (the so-called cumulative single star fraction).

not a function of the passband, the calibration field in LLN was also analysed with reference to Kp magnitudes, and the results are illustrated in Fig. 4.5. Here the single star fraction is at a minimum in the bin $15 < Kp < 16$, at 47.1%, and is 47.5% at $4 < Kp < 16$. The minimum occurs in the same place in Kp (Fig. 4.5) as in G (.Fig. 4.3). In Kp there are systems present in the bin $4 < Kp < 5$, both single star and binary, but the single star fraction is so low that to include the data in Fig. 4.5 would distort the figure. This is in contrast to Fig. 4.4, where there are no systems in the bin $4 < G < 5$. Similarly, the peak at $5 < Kp < 6$ in Fig. 4.5 is also down to small number statistics. Otherwise, the trends in the LLN field at $8 < Kp < 26$ are remarkably similar to those at $8 < G < 26$, in contrast to those in the Long Look South field at $8 < G < 26$.

The synthetic *Kepler* field from Farmer et al. (2013), matched to the targeting priorities of the original *Kepler* input catalogue, has 45% single stars where the limiting magnitude is $Kp \leq 16$. It is interesting is that the synthetic field from Farmer et al. (2013), created to match as far as possible the parameters of the original *Kepler* target list, has a lower proportion of single stars at $Kp \leq 16$ than the more general PLATO simulations.

The binary fraction as a function of magnitude deviates from the intended value of 0.5 in Fig. 4.3, Fig. 4.4 and Fig. 4.5, as well as in the field described in Farmer et al. (2013). It is higher when stars with $Kp < 16$ are considered, and lower when stars with $Kp \geq 16$ are considered. Nine sections of shell, three each at a distance of 0.5 kpc, 1.0 kpc and 2.0 kpc,

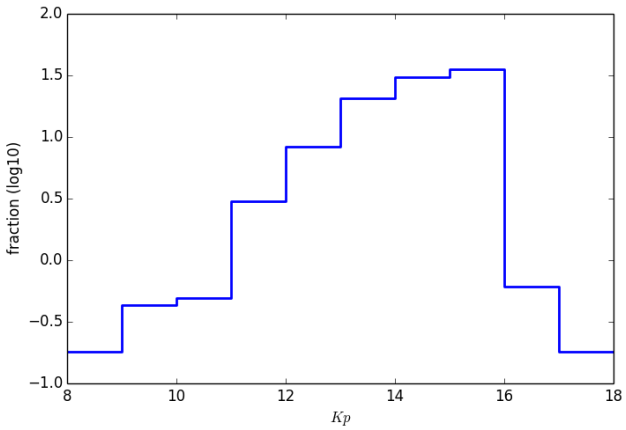


Fig. 4.6 Magnitude distribution of known *Kepler* planet hosts. Data from NExSci, *Kepler* DR25, accessed 21st July 2017.

were examined. These reveal the significance of the physically-motivated minimum stellar mass in `BiSEPS` simulations of 0.1 $M_{\odot}$. At 0.5 kpc and 1 kpc the minimum luminosity as a function of mass in unevolved stars can be observed. Clear trends indicate that low mass single stars are consistently fainter than binaries where the mass of the primary is equal to the mass of an equivalent single star. The binary fraction is ≥ 0.5 at $Kp \leq 16$, because binaries, which combine the luminosity of two stars, are intrinsically more luminous than single stars. As the intended binary fraction of 0.5 is essentially arbitrary, as described in Chapter 2; as blending with background binaries is unlikely to result in an observable eclipse where the difference in magnitude $\delta m \geq 10$; and as only the PLATO P5 population (red dwarfs) considers stars fainter than $V = 13$, the effect of this discrepancy on this work is small.

Fig. 4.5 indicates that binaries may outnumber single stars at the fainter magnitudes observable by *Kepler*, within the *Kepler* field. This raises the possibility that planets around stars with $15 < Kp < 16$ may be in unrecognised multiple systems, diluting the transit and leading to an estimate of the planet radius that is smaller than the true planet radius. $15 < Kp < 16$ is the region where the distribution of confirmed planet hosts peaks, as shown in Table. 4.2 and Fig. 4.6.

4.3 Summary: Chapter 4

This chapter has described the synthetic Galaxy used in my transiting exoplanet and eclipsing binary simulations. The regions simulated are described and compared with current thinking on the location of the two Long Look PLATO fields. The fraction of single stars as a function

Table 4.2 Magnitude distribution of the 1630 *Kepler* planet hosts. Data from NExSci, accessed 21st July 2017.

Kp	Number	Percentage
8.0 to 9.0	3	0.18
9.0 to 10.0	7	0.43
10.0 to 11.0	8	0.49
11.0 to 12.0	49	3.01
12.0 to 13.0	137	8.40
13.0 to 14.0	336	20.61
14.0 to 15.0	501	30.74
15.0 to 16.0	576	35.34
16.0 to 17.0	10	0.61
17.0 to 18.0	3	0.18

of magnitude in the synthetic Galaxy is evaluated both in the Northern and Southern fields used in this work, and real differences between the two fields are found. Concerns that the distribution of binaries as a function of mass and spectral type do not match observations are acknowledged, and addressing these issues is recommended for future work.

To determine the detectability by PLATO of the synthetic transiting exoplanets and eclipsing binaries simulated in this work, it is essential to understand the precision of the PLATO observatory. This, as well as concerns raised by PLATO's large plate scale, are examined in the next chapter.

Chapter 5

Detector Specifications

This chapter considers the factors that will influence PLATO's SNR (signal to noise ratio), as well as the impact of the large pixel scale on observations and target selection. Lessons that may be learned from instrumental effects experienced by Kepler will also be considered.

PLATO normal cameras will be situated in four groups of six cameras per group, which will increase the ability of PLATO to detect shallow signals as the data from the six cameras in each group will be combined to improve the SNR. The shallowest signals will be detected in the region where all four groups of cameras overlap, which may or may not be the most densely populated region of the on-sky field: see Fig. 1.5 for an illustration of how the overlap may work.

Signal and noise will be considered in the first section, and the pixel scale in the second section. A planet transit, or binary eclipse, is assumed to be detectable if the signal exceeds the noise by a suitably large factor. Specifically, the requirement that the transit SNR should exceed 7 is often used (Tenenbaum et al., 2012). Transit SNR is defined as

$$\mathrm{SNR}_{\mathrm{transit}} = \frac{\delta F \sqrt{N}}{\mathrm{noise}} \tag{5.1}$$

where δF is transit depth as defined in Eq. 1.1 in relation to planets and in Eq. 1.12 in relation to eclipsing binaries (the signal), and N is the number of observed events (transits or eclipses).

5.1 Signal and Noise

5.1.1 Signal

Signal represents the data of interest to the scientific community. In a transiting exoplanet survey utilising CCDs, this consists of photons, read out as photoelectrons. Variation in signal, for example during the transit of an exoplanet, is of particular interest.

This work has benefited from code provided by Claude Catala (Catala 2016, priv. comm.) enabling signal and noise to be calculated as a function of magnitude and effective temperature, and which has been adapted and incorporated into my own codes. This takes the throughput τ at 50 nm intervals in the range $500 < \lambda < 1000$ nm and interpolates at finer resolution. The throughput has changed during the instrument design process, and therefore the values provided by Catala, which date back to the M1/M2 application, are no longer valid. Contact was therefore made with Anko Börner, Head of Real-Time Data Processing, Institute of Optical Sensor Systems, German Aerospace Centre (DLR) (Deutsches Zentrum für Luft- und Raumfahrt), who kindly provided by private communication in July 2016 a copy of the PLATO Instrument Noise Budget dated 19 February 2016, which appears to be an internal document. Referring to this document, and supplementary tables provided by Dr Börner, updated throughput values have been calculated. τ is

$$\tau = \frac{n_{pe}}{n_{ph}} \tag{5.2}$$

where n_{pe} is the number of photoelectrons per pixel

$$n_{pe} = n_{pe,t} + n_{pe,c} + n_{pe,b} + n_{pe,s} \tag{5.3}$$

that is, the sum of photoelectrons from the target, contaminating photoelectrons, background photoelectrons and stray photoelectrons, terms which will be defined more fully later in this subsection, and n_{ph} is the number of photons per pixel

$$n_{ph} = n_{ph,t} + n_{ph,c} + n_{ph,b} + n_{ph,s} \tag{5.4}$$

again, the sum of photons from the target, contaminating photons, background photons and stray photons. So, in the formalism adopted in the Instrument Noise Budget, which I have followed, n_{ph} is the sum of the photons from all sources, and n_{pe} is the corresponding sum of photoelectrons. All sources of photons, and hence photoelectrons, which contribute to the signal are considered.

Photons

The number of photons from the target,

$$n_{ph,t} = A_{ap} t_{exp} \sum_{i=1}^{N} D_{ph}(i) \tag{5.5}$$

depends on the photon density D_{ph} in ph cm^{-2} s^{-1}, which is itself a function of the target's magnitude; the area of the aperture A_{ap} which is $12^2\pi/4$ cm^2, derived from the pupil diameter of 12.0 cm and so is the area of a circle with a diameter of 12.0 cm; and the exposure time, t_{exp}, which is 22 s. i represents bins of 50 nm width and N is the number of bins. Although it is anticipated that the optics' spectral range will be 500 nm to 1000 nm (PLATO Definition Study Report), the simulations reported in the Instrument Noise Budget use a broader passband of 400 nm to 1100 nm: so here N is 15. The code provided by Catala covers the wavelengths 500 nm to 1000 nm, the same as the optics spectral range reported in the PLATO Definition Study Report, so here N is 11.

Tables are provided in the Instrument Noise Budget for the photon density in ph cm^{-2} s^{-1} at $V = 0$, 8 and 11, at the appropriate bin width of 50 nm.

The number of background photons, $n_{ph,b}$, is derived from the number of background photoelectrons, $n_{pe,b}$, which is set at a constant, 100 pe px^{-1} s^{-1}. This estimate is referenced to PLATO-INAF-PL-RP-003 i3.1, the Stray Light Analysis Report, also an internal document. Background photons come from zodiacal light and stray light from the Sun and Moon. As the sum of these sources of background photons is likely to remain constant over the course of the mission, it is appropriate to use a constant to describe them.

The number of contaminating photons, $n_{ph,c}$, is arbitrarily set at 1% of the number of photons from a $V = 11$ star.

The number of photons from stray light, $n_{ph,s}$, is derived from the number of stray photoelectrons, $n_{pe,s}$, which has two contributions: 2 pe px^{-1} s^{-1} from the spacecraft, and 20 pe px^{-1} s^{-1} from the instrument optics. These estimates are referenced to PLATO-DLR-PL-RS-001 i3.1, Instrument Technical Requirements Document, also an internal document.

$n_{pe,b}$ and $n_{pe,s}$ are given as constants in units of pe px^{-1} s^{-1}, but, as described in the following subsection, because the conversion between photoelectrons and photons depends on a miscellaneous efficiency factor, E, which is a function of magnitude, $n_{ph,b}$ and $n_{ph,s}$ are themselves magnitude dependent. Since it is the number of photoelectrons rather than the number of photons which is recorded and used for further analysis, following readout of the CCDs, this approach allows the observer to discount the first $\approx$ 122 pe px^{-1} s^{-1} in any data as due to background and stray sources and is certainly convenient for follow up work.

However, it does mean that care needs to be taken when converting from photoelectrons to photons to account for the magnitude dependence of E.

Photoelectrons

The number of photoelectrons, n_{pe}, is the sum of the photoelectrons from the target, $n_{pe,t}$, the background, $n_{pe,b}$, the contaminants, $n_{pe,c}$ and stray photoelectrons, $n_{pe,s}$.

$n_{pe,t}$ can be calculated by integrating the transmitted photon flux:

$$n_{pe,t} = A_{ap} t_{exp} E f_{sp} \int_{\lambda 1}^{\lambda 2} D_{ph,spec}(\lambda) T_{op}(\lambda) T_{fil}(\lambda) QE(\lambda) \mathrm{d}\lambda \tag{5.6}$$

which, after spectral discretisation, becomes

$$n_{pe,t} = A_{ap} t_{exp} E f_{sp} \sum_{i=1}^{N} D_{ph,spec}(i) T_{op}(i) T_{fil}(i) QE(i) \tag{5.7}$$

Here E is a miscellaneous efficiency factor which is weakly dependent on source magnitude, and is 0.870 for $m_V = 0$, 0.868 for $m_V = 8$ and 0.859 for $m_V = 11$. Efficiencies included in the miscellaneous term are those caused by vignetting (0.945), polarisation (0.989), particle contamination (0.98), molecular contamination (0.9566), the angular dependencies of the quantum efficiency (0.993) and charge transfer inefficiency (1 at $V = 0$, 0.998 at $V = 8$ and 0.987 at $V = 11$).

f_{sp} is the signal spread factor, arising from broadening of the signal due to the point spread function (PSF) and is set at 90% of the energy enclosed in 2.5 x 2.5 pixels, or 0.144 pe^{-1}.

$D_{ph,spec}(\lambda)$ is the spectral photon density from the star, measured in units of photons cm^{-2} s^{-1} nm^{-1}.

$T_{op}(\lambda)$ is the wavelength dependent transmissivity of the optics, and $T_{fil}(\lambda)$ is the transmissivity of the filters and applies to the fast cameras only. $QE(\lambda)$ is the quantum efficiency of the CCD. Values for $T_{op}(\lambda)$ and $QE(\lambda)$ are given in Table 5.1. Beginning of life rather than end of life figures are given in the table for $T_{op}(\lambda)$: end of life are slightly lower. The filter transmissivity, $T_{fil}(\lambda)$ is 0.85 for < 700 nm and 0 for > 700 nm in the blue filter, while it is 0 for < 700 nm and 0.85 for > 700 nm in the red filter. So when both filters are considered, $T_{fil}(\lambda)$ is constant at 0.85 over the whole spectral range.

The following equations describing the conversion from photons to photoelectrons presented in the Instrument Noise Budget for background, contaminating and stray sources are:

$$n_{pe,b} = D_{ph,b} t_{exp} E < T_{op} T_{fil} QE > \tag{5.8}$$

Table 5.1 Transmissivity of the PLATO optics and quantum efficiency of the CCDs as a function of λ. Data from the PLATO Instrument Noise Budget and from a private communication from Dr Anko Börner.

λ/nm	$T_{op}(\lambda)$	$QE(\lambda)$
500	0.7332	0.7370
550	0.7640	0.8070
600	0.7713	0.8770
650	0.7769	0.8675
700	0.7842	0.8580
750	0.7899	0.7565
800	0.7939	0.6550
850	0.7955	0.4845
900	0.7972	0.3140
950	0.7980	0.165
1000	0.7996	0.062

$$n_{pe,c} = D_{ph,c} t_{exp} E f_{sp} < T_{op} T_{fil} QE > \tag{5.9}$$

$$n_{pe,s} = D_{ph,s} t_{exp} E < T_{op} T_{fil} QE > \tag{5.10}$$

$D_{ph,b}$ is the background photon density, $D_{ph,c}$ is the photon density from contaminants and $D_{ph,s}$ is the stray photon density, $< T_{op} T_{fil} QE >$ is a suitable wavelength average of transmissions and quantum efficiencies. All other terms have been previously defined.

Throughput τ

Using the formulae and the data from the Instrument Noise Budget, the throughput was recalculated for this work as a function of wavelength. Throughput is, of course, a constant for a given instrument. An estimate of throughput was not explicitly included in the PLATO Instrument Noise Budget of 2016. However, all the tools to obtain an estimate were, in the form of equations and quantities at the three magnitudes: V = 0, 8 and 11. I therefore considered three estimates of throughput, based on the three sets of parameters in the Instrument Noise Budget. These estimates consider all photons incident on the camera and all counted photoelectrons, regardless of source. I term these "effective throughput" to emphasise the point that throughput is a constant and is not magnitude dependent, while these estimates are. I then evaluated the three estimates of effective throughput to determine which, when incorporated into Claude Catala's code, provided the closest match to the photometric

Table 5.2 PLATO effective throughput, estimated using expressions from the PLATO Instrument Noise Budget. The estimate based on $V = 11$ was found to be the best estimate for a source independent throughput. See the text for more details.

λ/nm	Previous value $V = 0$	Updated value $V = 0$	Updated value $V = 8$	Updated value $V = 11$	Ratio $\frac{V=11}{V=0}$
500	0.627	0.360	0.359	0.353	0.980
550	0.739	0.410	0.409	0.397	0.968
600	0.780	0.450	0.448	0.432	0.960
650	0.799	0.449	0.447	0.430	0.958
700	0.785	0.448	0.446	0.429	0.958
700	0.731	0.398	0.396	0.386	0.970
800	0.628	0.346	0.345	0.341	0.986
850	0.486	0.257	0.257	0.265	1.031
900	0.325	0.167	0.168	0.191	1.144
950	0.174	0.088	0.090	0.126	1.432
1000	0.066	0.032	0.037	0.084	2.625

precision described in the PLATO Instrument Noise Budget, and use that single estimated effective throughput in my simulations.

It was found that the throughput had significantly reduced since Catala's code was originally written. A comparison is given in Table 5.2 and Fig. 5.1 for the three effective throughputs calculated in this work as a function of magnitude: $V = 0$, 8 and 11. The calculated figures for $V = 0$ and $V = 8$ are very similar, while $V = 11$ diverges, especially in the infrared. This is due to the fact that the number of photons expected to be received from a $V = 11$ star is lower than in the other cases, so such a star will be more influenced by background, contaminating and stray sources than brighter stars.

The new throughput is significantly lower than that derived by Catala for the M1/M2 bid in 2009 (as acknowledged in the caption to Fig. 5.4, Rauer et al. (2014)). Marcos-Arenal et al. (2014), dating from the time of the successful M3 bid, describes a PLATO simulator. Comparison of Eq. 3 in Marcos-Arenal et al. (2014) with the equations describing the photoelectrons in the PLATO Instrument Noise Budget (19 February 2016) indicates that the significantly lower throughput in this work lies in the inclusion in the PLATO Instrument Noise Budget of the miscellaneous efficiency term E along with more detailed modelling of the transmission efficiency.

Table 5.3 compares the contribution of background, contaminating and stray photons and photoelectrons to the total number of photons and photoelectrons, as well as considering the ratios of photons and of photoelectrons at different wavelengths.

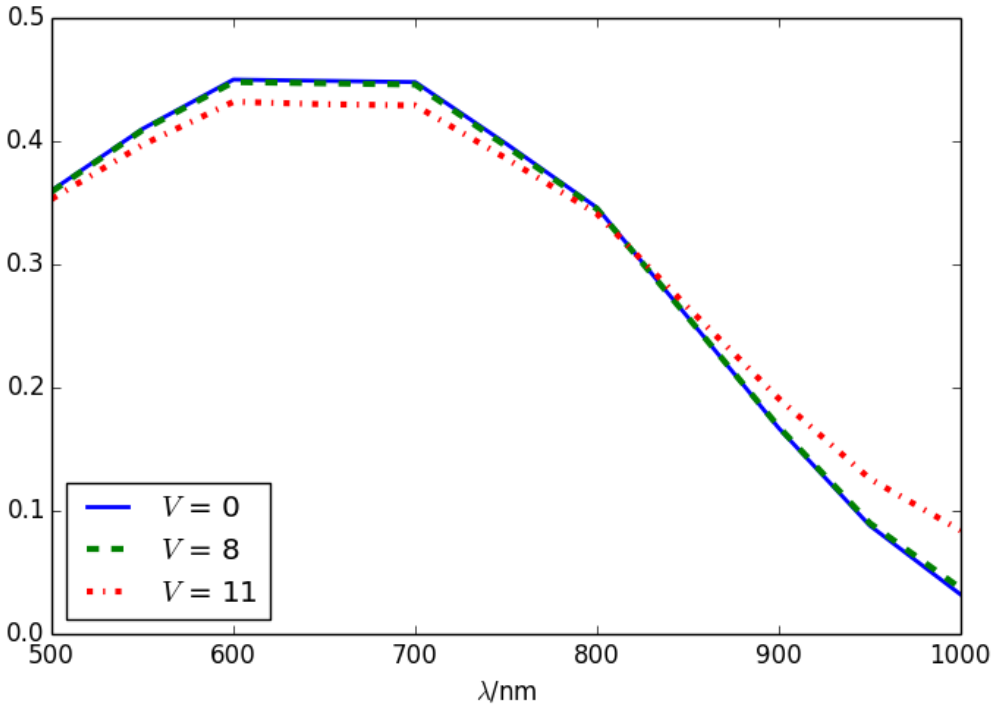


Fig. 5.1 Estimates of PLATO effective throughput derived in this work, using the PLATO Instrument Noise Budget to estimate the numbers of photons and photoelectrons at $V = 0$, 8 and 11 magnitudes. These estimates are used to identify single best source independent throughput. See the text for more details.

Table 5.3 Comparison of contributions of stray, background and contaminating sources to the PLATO signal at $V = 0$, $V = 8$ and $V = 11$.

Ratio	λ	$V = 0$	$V = 8$	$V = 11$
$(n_{pe,b} + n_{pe,c} + n_{pe,s})/n_{pe}$	600 nm	0.0005%	0.73%	10%
$(n_{pe,b} + n_{pe,c} + n_{pe,s})/n_{pe}$	1000 nm	0.0083%	12%	68%
$(n_{ph,b} + n_{ph,c} + n_{ph,s})/n_{ph}$	600 nm	0.00059%	0.93%	13%
$(n_{ph,b} + n_{ph,c} + n_{ph,s})/n_{ph}$	1000 nm	0.00084%	1.3%	18%
$n_{ph,c}/n_{ph,t}$	600 nm	0.00004%	0.063%	1.0%
$n_{ph,c}/n_{ph,t}$	1000 nm	0.00004%	0.063%	1.0%
$(n_{pe,b} + n_{pe,s})/n_{pe,t}$	600 nm	0.00042%	0.67%	11%
$(n_{pe,b} + n_{pe,s})/n_{pe,t}$	1000 nm	0.0083%	13%	210%
n_{ph}/n_{ph}	600/1000 nm	1.44:1	1.44:1	1.37:1
n_{pe}/n_{pe}	600/1000 nm	19.7:1	17.5:1	7.04:1
$n_{pe,t}/n_{pe,t}$	600/1000 nm	19.7:1	19.7:1	19.7:1

Table 5.4 Comparison of n_{pe}/n_{ph} as given in the PLATO Instrument Noise Budget and as calculated for this work, $V = 8$ and $V = 11$. BoL: Beginning of life, and EoL: End of life, both from the Instrument Noise Budget. Lifetime is expected to be approximately 6 years.

source	m_V	n_{pe}	n_{pe} px^{-1}	n_{ph}	n_{ph} px^{-1}	n_{pe}/n_{ph}
BoL	8	3328915	486841	10036446		0.331
EoL	8	3205509	468893	10036446		0.319
This work	8		411943		1266124	0.309
BoL	11	207725	32896	633257		0.328
EoL	11	183511	29374	633257		0.290
This work	11		29931		92377	0.324

In the infrared for a $V = 11$ star, photoelectrons from background, contaminating and stray sources can be more significant than photoelectrons from the target. This affects the numerator when calculating throughput as a function of magnitude. That this does not affect photons in the same way, and hence the denominator when calculating throughput as a function of magnitude, is likely due to factors which affect the number of photoelectrons but not the photons and which are wavelength dependent: transmissivity, which improves at longer wavelengths and quantum efficiency, which declines at longer wavelengths. Indeed, reference to Table 5.1 and Fig. 5.1 indicates the shape of the curve in Fig. 5.1 is strongly dependent on the quantum efficiency as a function of wavelength.

To verify the simplified calculation made here, reference was made to a table within the Instrument Noise Budget which gives the estimated total number of photons and photoelectrons for a star of a given magnitude, $8 < V < 11$, both at the beginning of life and end of life (approx 6 years after launch), integrated over $400 < \lambda < 1100$ nm. The results are given in Table 5.4.

As shown in Table 5.2, the original Catala code used $V = 0$, which in the new calculations has similar calculated values to the calculated values for $V = 8$. However, as shown in Table 5.4, there is divergence at $V = 8$ between the values in the Instrument Noise Budget, and those calculated here: specifically, for n_{pe}/n_{ph} and n_{pe} px^{-1}, the values calculated in this work are lower than the end of life (EoL) values for $V = 8$ (this work is 85% of BoL and 88% of EoL values, n_{pe} px^{-1}, and 93% of BoL and 97% of EoL, n_{pe}/n_{ph}). For $V = 11$, however, the values calculated in this work fall between the beginning of life (BoL) and end of life (EoL) values (this work is 91% of BoL and 102% of EoL values, n_{pe} px^{-1}, and 99% of BoL and 112% of EoL, n_{pe}/n_{ph}).

As the calculated ratio of photoelectrons to photons and the number of photoelectrons per pixel fall below what is expected during PLATO's six-year lifetime for $V = 8$, but for $V = 11$

are consistent with the expected ratio during the mission, the values for $V = 11$ are used for throughput in this work.

5.1.2 Noise

With any system there is noise. In the context of PLATO light curves, noise is defined as the rms value of light curves binned to one hour, where one hour is the integration time used in pre mission modelling. The Instrument Noise Budget describes the sophisticated modelling that takes place with PLATO, dividing the sources of noise into three different categories: line of sight related noise; instrument related noise; and system related noise (Table 5.5).

Levels of noise associated with the instrument and system may be expected to be updated as the instrument design progresses. Levels of noise associated with the line of sight will probably remain the same.

The two most significant sources of instrument related noise in the Instrument Noise Budget are front-end electronic (FEE) readout noise, which is caused by the output amplifier on the front end electronics, and CCD readout noise, which by caused by the output amplifier on the CCD.

The noise calculation in the Catala code is obviously much simpler than the full model presented in the Instrument Noise Budget, but where corresponding values are available in both sources, they do appear to be in line.

With the original throughput estimate, the estimate of noise obtained by the Catala code is illustrated in Fig. 5.4 of Rauer et al. (2014) (noise is the term used on the y axis of this figure).

As well as the throughput, the number of cameras per group has also been amended since Rauer et al. (2014). The original proposal was for eight normal cameras/group plus two fast cameras, and final approval for PLATO was based on six normal cameras per group plus two fast cameras. The effect of the change in the number of cameras is illustrated by comparing Fig. 5.2 (32 cameras) and Fig. 5.3 (24 cameras), and the effect of the change in throughput and in the number of cameras is indicated in Table 5.6.

Table 5.6, Fig. 5.2 and Fig. 5.3 all make reference to noise levels of 34 ppm in one hour and 80 ppm in one hour (where 'noise' is the term used by the Catala code). These were specified in the science requirements for the PLATO mission. 34 ppm is the target for the P1 and P2 in the central part of the field, where all four cameras are observing, as this is necessary to carry out asteroseismology on the stars, whether they are planet hosts or not. As described in Chapter 1, this will enable the constraints on the mass, radius and age of the stars observed to be significantly improved. 80 ppm is what would be required to for an Earth-like planet orbiting a Sun-like star to produce a measurable transit (although whether

Table 5.5 PLATO sources of noise. Column 1: type of noise. Column 2: value in the Instrument Noise Budget (INB), camera level. Column 3: value in the INB, instrument level. "Jitter" refers to instances where the image is moving around on the CCD, and the associated noise includes masking, photo response non-uniformity and dark signal non-uniformity. "Miscellaneous photon noise" includes background, contaminating and stray photons, already described in the Section 5.1.1.

source	INB, camera	INB, instrument
Line of sight		
Jitter	108 ppm	9 ppm
Aberration	n/a	< 0.1 ppm
Image scale and distortion	1828.97 ppm	0.27 ppm
Instrument related noise		
Optics, including:		
PSF blurring	< 0.1 ppm	< 0.1 ppm
PSF breathing	10 ppm	0.15 ppm
Pupil variation noise	4 ppm	< 0.1 ppm
Miscellaneous photon noise	54.62 ppm	2.02 ppm
CCD related noise, including		
Quantum efficiency	18.5 ppm	0.15 ppm
Dark current	104.47 ppm	0.13 ppm
Dark current shot	183.11 ppm	0.97 ppm
CCD readout	1480.99 ppm	7.89 ppm
CCD gain stability	55.56 ppm	0.46 ppm
CCD smearing photon	388.44 ppm	2.07 ppm
Front end electronic noise		
FEE gain stability	16.67 ppm	< 0.1 ppm
FEE offset stability	111.44 ppm	< 0.1 ppm
FEE readout	1353.95 ppm	7.21 ppm
ADC quantisation	193.01 ppm	1.03 ppm
CCD bias voltage stability	34 ppm	0.71 ppm
System related noise		
Oscillator stability	0.002 ppm	0.36 ppm
EMC noise	TBD	TBD
Processing noise	TBD	TBD

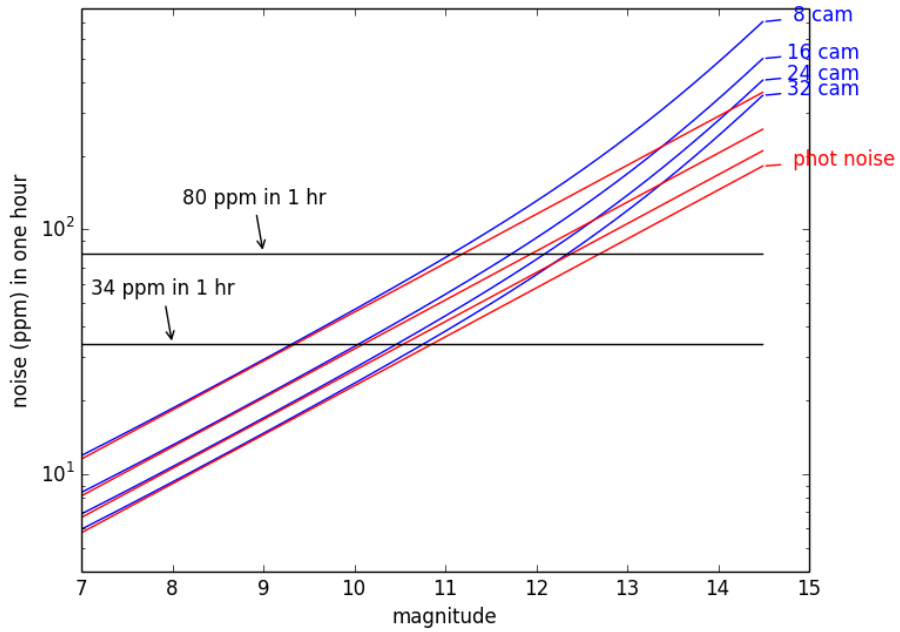


Fig. 5.2 Predicted minimum detectable signal for PLATO as a function of magnitude, 32 cameras (8 cameras/group), using the new throughput calculated for this work and a one hour integration.

Table 5.6 Magnitude limit where noise is 34 ppm and 80 ppm in 1 hour, by number of cameras per group and by throughput.

Type of throughfeed	Noise /ppm	Cameras /group	1 group *V*	2 groups *V*	3 groups *V*	4 groups *V*
Original	34	8	9.95	10.67	11.08	11.37
Revised	34	8	9.31	10.04	10.46	10.75
Original	34	7	9.80	10.53	10.95	11.24
Revised	34	7	9.17	9.90	10.32	10.62
Original	34	6	9.64	10.37	10.79	11.08
Revised	34	6	9.00	9.74	10.16	10.46
Original	80	8	11.69	12.33	12.68	12.92
Revised	80	8	11.08	11.74	12.10	12.35
Original	80	7	11.56	12.21	12.57	12.81
Revised	80	7	10.94	11.61	11.98	12.24
Original	80	6	11.41	12.07	12.43	12.68
Revised	80	6	10.79	11.47	11.84	12.10

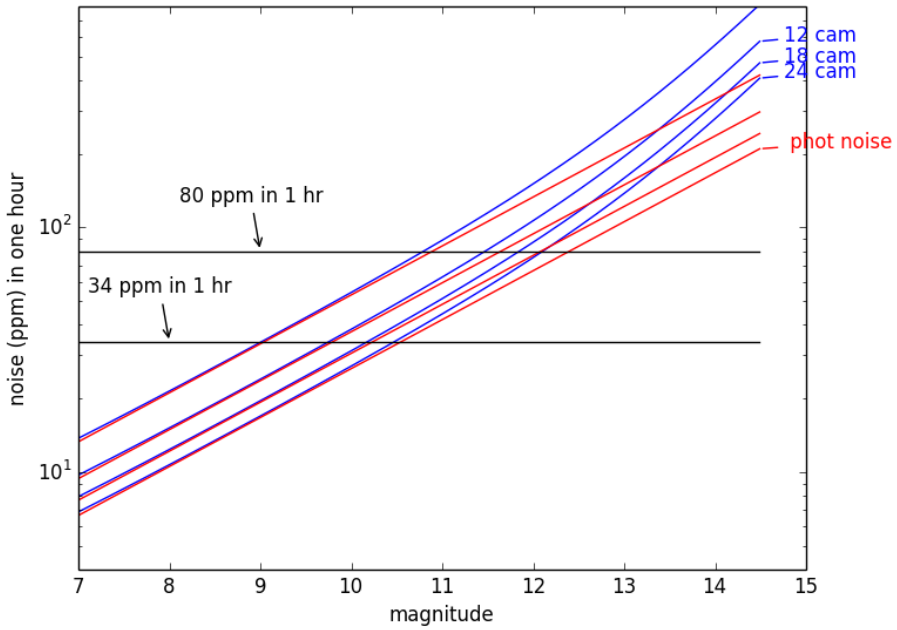


Fig. 5.3 Predicted minimum detectable signal for PLATO as a function of magnitude, 24 cameras (6 cameras/group), using the new throughput calculated for this work and a one hour integration.

it was detectable would depend on the transit SNR, as defined in Eq. 5.1) when appropriate values are inserted into a rearranged version of Eq 1.1:

$$\delta F = \left(\frac{R_P}{R_*}\right)^2 \tag{5.11}$$

Such a planet is a priority for PLATO. 80 ppm in a one hour integration is the target for the central part of the field for the P5 population. The target for the P4 population is 800 ppm in a one hour integration in the central part of the field. For easy reference, the nature of the stars in each of the PLATO priority populations, P1, P2, P4 and P5, is described in Table 1.6.

Table 5.6 indicates that the reduction in the number of cameras per group from eight to six does reduce the magnitude at which the transit of an Earth-like planet can be observed, but more significant is the change in throughput. The data in Table 5.6 indicates that targets may be missed for the fainter members of P1 and P5 in the central field. Fig. 5.2 and Fig. 5.3 were generated using a model star with an effective temperature of 6000 K.

It should also be noted that the model used in this work is far less sophisticated than the models described in the Instrument Noise Budget. However, their Table 6.9 does carry out the exercise of indicating photon noise, random noise at an instrument level and total noise for a $V = 11$ star for a one hour integration, for 32, 28 and 24 cameras, both at BoL and EoL. These

figures are consistent with the results presented in Fig. 5.2 and Fig. 5.3 for all four groups of cameras in a one hour integration. For 32 cameras, the estimate in the Instrument Noise Budget for photon noise at an instrument level is 32.3 ppm at BoL, 34.4 ppm at EoL (the equivalent of the orange line on the figure); and total noise (the equivalent of the blue line on the figure), calculated by taking the square root of the sum of the squared photon noise and random noise of the instrument, is 34.2 ppm at BoL, 36.6 ppm at EoL. Similarly, for 24 cameras, photon noise at an instrument level is 37.3 ppm at BoL, 39.7 ppm at EoL; while the total noise is 39.5 ppm at BoL, 42.2 ppm at EoL. The integration time is one hour.

The consistency between our results and those from the Instrument Noise Budget indicate that, pending changes in instrument design, the results generated by the updated Catala code can be trusted in determining the noise level for a star of a given temperature and magnitude, for a given number of groups of cameras and for a given integration time.

Overlap

A second code provided by Claude Catala, also integrated into our work, allows the overlap of the four cameras to be calculated. This is illustrated in Fig. 1.5. Given a specific reference point as the centre of the field, for example $l = 65°$, $b = 30°$, it is therefore possible to determine how many cameras will be able to observe a given star and, given its magnitude, it is possible to determine whether or not a planet of a given radius orbiting a given host star at a given position would be observable, assuming the planet is transiting.

Other projections of the overlap do exist, and Fig. 5.4, from the PLATO Definition Study Report, uses an overlap pattern with straight lines instead of curves, and with sharper corners. However, the central region, in which all four camera groups overlap appears similar in proportion to the total field in both Fig. 5.4 and Fig. 1.5. This study does not have access to the code which generated the overlap shown in Fig. 5.4, therefore the alternative overlap, illustrated in Fig. 1.5, is used.

A simplifying assumption is that the boundaries of the synthetic fields approximately follow lines of Galactic longitude and latitude. This is because the precise fields are still under consideration. If the level of contamination is favourable, it may be possible to adjust the angle of the field of view so that, while the centre of the field of view remains within the region indicated in Fig. 5.4, it is possible to observe closer to the Galactic plane. This will depend on the proportion of possible planets to contaminants in the relevant region, and consideration of this is beyond the scope of this work.

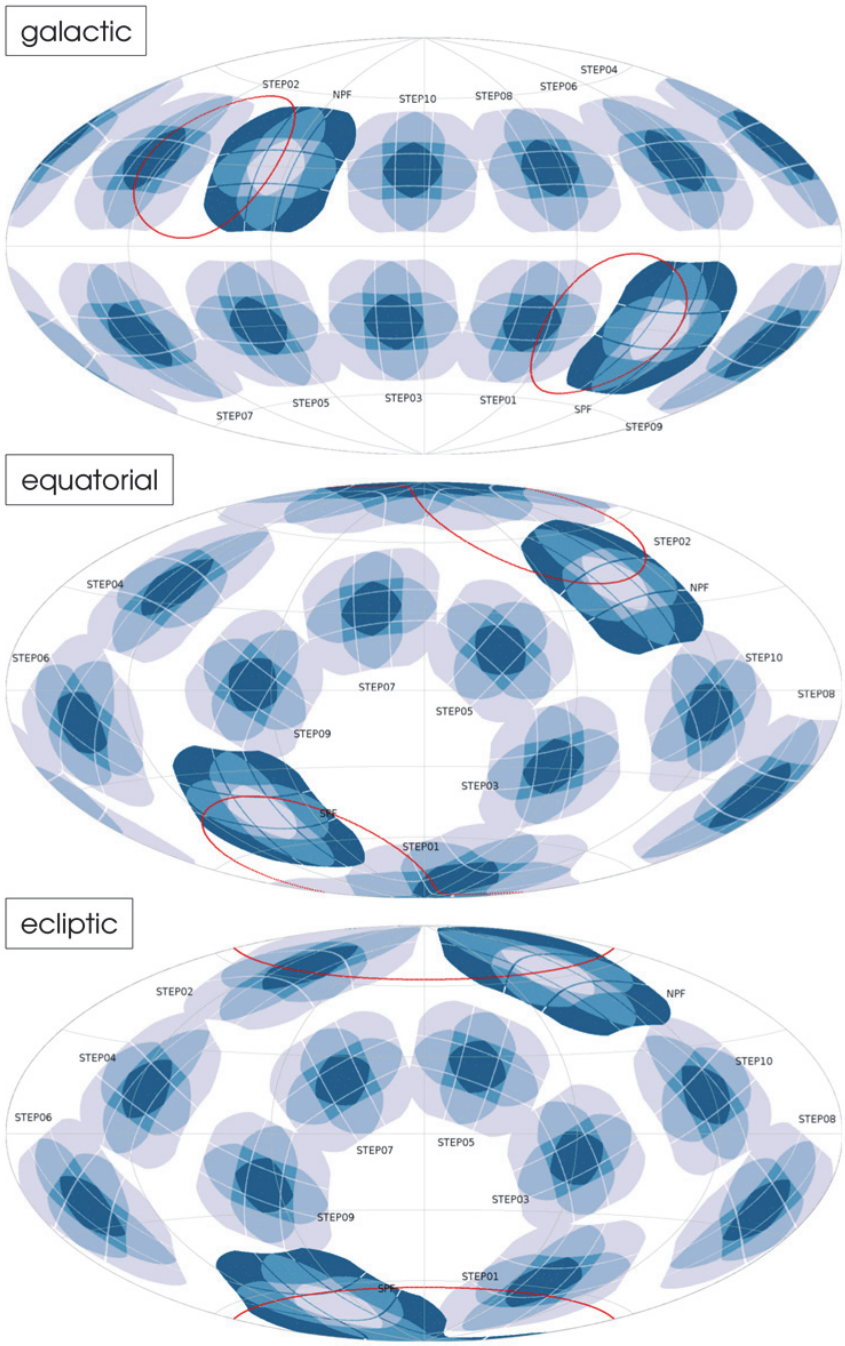


Fig. 5.4 Aitoff projections, from the PLATO Definition Study Report, of the two preliminary long look fields (SPF, NPF) and ten step-and-stare fields (STEP01– STEP10), all centred at |b| = 30°, in the Galactic (upper plot), equatorial (middle), and ecliptic (lower) reference frames. The red lines enclose the long look field pointing requirement limits. The long look fields are color-coded on an inverted scale.

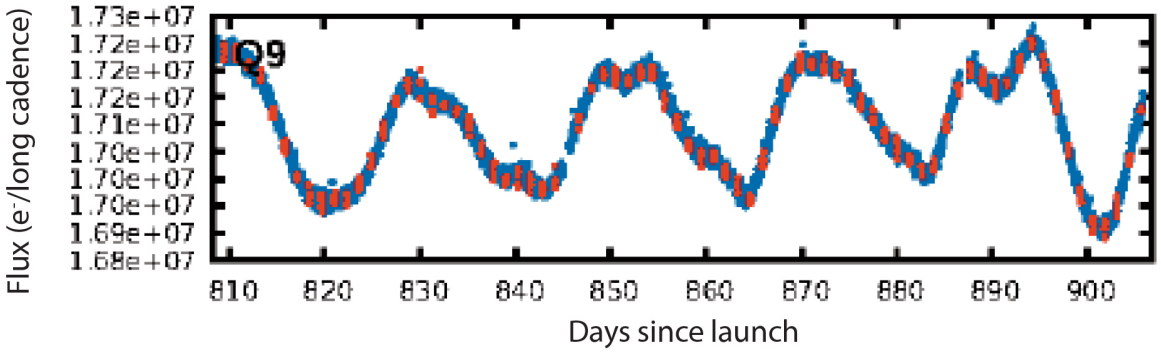


Fig. 5.5 Starspots in *Kepler* data: PDC time series, KOI 4844.01. The signal from a $0.87^{+0.13}_{-0.08}$ $R_\oplus$ Kepler planet candidate, with a period of 1.28 days, a transit depth of 126 $\pm$14.1 ppm and a transit duration of 1.330 hours (red) is far less significant than the signal from starspots. Q9 selected at random to demonstrate this effect. The stellar parameters for the host are $T_{eff} = 5285^{+159}_{-143}$ K (source: photometry), $\log g = 4.566^{+0.049}_{-0.091}$ (source: KIC), $R = 0.775^{+0.121}_{-0.069}$ $R_\odot$, $M = 0.805^{+0.086}_{-0.078}$ $M_\odot$ (source: Dartmouth stellar models). All data from NExScI.

Other considerations

Stellar variability means that the flux from a star is not a constant. In particular, if a star is active, a transiting planet may be more difficult to detect than the figures given earlier in this chapter may suggest. Activity is more likely in stars earlier than F5, and in M class stars. Hence, the concentration by PLATO on F5-K class stars in the majority of the priority classes.

Software is likely to be used to search for repeating patterns that may indicate a transiting planet, as with other transiting exoplanet surveys. These can be very shallow, especially compared to the noise from natural stellar variability, as shown in Fig. 5.5, where the signal from an 0.87 $R_\oplus$ planet candidate (KOI 4884.01) at 126 ppm and P = 1.28 days is far less significant than the $\approx$ 20 day signals associated with starspots.

In Fig. 1.7 the dip indicating a highly grazing eclipsing binary in KIC 9027841 would be difficult to identify without such software. The very short orbital period of KIC 9027841 of 0.504 days is useful in this regard, as in an 80 day pointing it would have been observed 158 or 159 times. The eclipse depth is 975.6 $\pm$6.9 ppm. *Kepler*-10 c, with diagnostic information illustrated in Fig. 1.8, is a 2.26 $R_\oplus$ planet with an eclipse depth of 447.1 $\pm$3.7 ppm, orbiting a host star with an estimated effective temperature of 5676 K and an estimated radius of 1.044 $R_\odot$. As shown in Fig. 1.8, the transit is shallow enough that the shoulders just before and just after transit are visible in the light curve and, again, natural stellar noise is clearly visible. With an orbital period of 45.29 days this planet would have been observed twice, possibly once, in an 80 day pointing.

As demonstrated by Eq. 5.1, phase folding allows shallow, short period threshold crossing events such as KOI 4884.01 (Fig. 5.5) and KIC 9027841 to be identified from data, by significantly improving the SNR. This is covered in more detail in Appendix A of Tenenbaum et al. (2012), but essentially a multiple event statistic (MES), l, is found using

$$l \equiv \frac{\sum_k \mathbb{N}(k)}{\sqrt{\sum_k \mathbb{D}(k)}} \tag{5.12}$$

where k runs over the locations of the prospective transits, $\mathbb{N}$ is the sum of the contributions at all frequencies to the detection of a particular signal in a flux time series, and $\mathbb{D}$ is the expected detection amplitude of the noise spectrum. If values of $\mathbb{N}$ and $\mathbb{D}$ remain consistent in multiple prospective transits, the detection significance will rise as the square root of the number of events. Tenenbaum et al. (2012) reported that a MES of 7.1 is considered significant in *Kepler* analysis as it reduces the number of false positives while at the same time not eliminating Earth-like events.

Lessons applicable from *Kepler* on further instrumental effects

Recent work by the *Kepler* team, described in Coughlin (2017), Thompson et al. (2017) and Bryson et al. (2017), investigated the number of false positives due to instrumental effects, as well as the reliability of the *Kepler* pipeline. This work may be applicable to PLATO in identifying the type of instrumental effects that may occur. Naturally, as PLATO is a different instrument it will experience unique instrumental effects. The methods of identifying and eliminating these effects which may mimic the transit of an Earth-like planet in *Kepler* may well be informative, however.

Three experiments were carried out by the *Kepler* team. In one, transits of both planets and eclipsing binaries (blended and unblended) were injected into the *Kepler* pipeline, and the proportion successfully recovered and correctly identified was examined. This is described in more detail in documentation on the NExSci website[1]. This the *Kepler* team regard as confirming the relative completeness of the DR25 catalogue. This was the fourth injection experiment carried out by this team, and is a useful way of verifying the accuracy of the transit detection software.

Reliability was investigated in two additional experiments, which involved inverting and scrambling genuine transits, in an attempt to simulate periodic and non periodic false alarms respectively. A false alarm differs from a false positive in that a false alarm arises from instrumental effects or stellar variability while a false positive represents an eclipsing binary and/or a centroid offset. The resulting files are available on the NExSci website for download[2].

[1]https://exoplanetarchive.ipac.caltech.edu/docs/KSCI-19110-001.pdf
[2]https://exoplanetarchive.ipac.caltech.edu/docs/KeplerSimulated.html

An example of a periodic false alarm can be found in the excess of Threshold Crossing Events (TCEs) around 372 days: 372 days is the *Kepler* orbital period, and it is believed that thermally coupled events are behind this excess. A particular feature of these instrumental effects was that there would be a sudden increase as well as a sudden decrease in signal, producing a quasi-sinusoidal effect: software looking for only a decrease in flux would therefore select these for further study, disregarding the fact that there had also been an increase in signal. Inverting the signal means such quasi sinusoidal events are still detected, while genuine planets are not as they only show an increase in flux when inverted. This enables the rate of such thermally coupled events to be quantified.

Examples of non-periodic false alarms are pixel drop-outs and statistical fluctuations which randomly line up to give the impression of a transit. This will particularly affect long period planet candidates. When the quarters of the dataset are scrambled, genuine planets will no longer be detectable but pixel drop-outs and statistical fluctuations will remain.

Both the experiments in which data was inverted and scrambled were successful in recovering false alarms. A sharp peak in the TCEs around 372 days was successfully recovered by the inverted signal experiment, while a broader increase in TCEs also centred on 372 days was successfully recovered by the scrambling experiment. These experiments indicate that *Kepler* TCEs with a year-long period should be treated with extreme caution.

Pre flight testing on *Kepler* indicated another source of false alarms, the so-called "rolling bands", a time-varying moiré pattern affecting certain channels (Kolodziejczak et al., 2010). Two principle sources were found: crosstalk between the 84 science CCDs and the four fine guidance sensor CCDs, and a high frequency amplifier oscillation in $< 40\%$ of the CCD readout channels. As described in Kolodziejczak et al. (2010), methods were found to flag suspect data, which was found to affect only a small fraction of the array at any given time.

PLATO designers will undoubtedly have taken advantage of *Kepler* publications on this issue to ensure that, as far as possible, PLATO is not affected by rolling bands. However, unanticipated instrumental effects unique to PLATO may emerge in the instrument testing phase, which cannot be accounted for at this point in time.

PLATO will orbit in the L2 position, while *Kepler* is Earth-trailing: but PLATO will also be tied to the Earth's orbital period, so is as likely to be affected by instrumental effects which mimic shallow transits with an orbit of about one year as *Kepler* is, although the particular effects may be different. PLATO's 24 cameras in four groups of six increase the probability that a sudden pixel drop out in one camera will not be replicated in others, reducing the risk of this particular type of false alarm, for example. However, given that a particular aim of PLATO is to detect an Earth-twin, with a year-long orbit, this research from the *Kepler* team suggests that careful visual inspection of such signals to check for instrumental effects before ground follow up would be a wise precaution.

5.2 Pixel Scale

The large pixel scale of 15.0" is one of the primary causes of concern regarding false positives from astrophysical sources in the PLATO mission. Further, as shown by the value for the signal spread factor, f_{sp} above, which considers the PSF over 2.5 x 2.5 pixels, it is anticipated that flux from a stellar system will affect more than one pixel. The Instrument Noise Budget states that 90% of the energy will be included in 2.5 x 2.5 pixels. The PLATO Definition Study states that the PSF surface will always be included within 9 pixels and that imagettes of 6 x 6 pixels (normal cameras) and 9 x 9 pixels (fast cameras) will be obtained for each target star, with some raw data being sent to the ground for processing and the remaining raw data being processed on board, to reduce the volume of telemetry data.

Kepler, with a 3.98" pixel size, found that pixels would saturate at about $Kp = 11$,[3] where one exposure is taken every 6.5 s[4] and integrated over either 1765.5 s (29.4 min) cadence or 58.89 s cadence. Charge from saturated pixels could bleed along columns, affecting analysis of nearby stars: and obviously, the brighter the star in the Kp pass band, the greater the problem of charge bleed. PLATO is specifically designed to target bright stars, so such issues are being taken into account during the instrument design process.

5.2.1 Case study from *Kepler*: KOI 102.02

A good example from *Kepler* of the problems that may arise through contamination by bright nearby stars is KOI 102 (KIC 8456679). This system has two TCEs: .01 is a planet candidate, with an estimated radius of $3.27^{+0.54}_{-0.33}$ $R_\oplus$ and a period of 1.735 days, and the other, .02, is listed as a false positive (on the grounds of a centroid offset, indicating that the signal may not be centred on the observed star): the period is 4.068 days and the depth would indicate a planet of $0.98^{+0.17}_{-0.1}$ $R_\oplus$, if it is genuine. And it may well be: new for DR25 is a score which shows how likely a TCE is to be a planet or a false positive, with 1.0 representing a TCE that the pipeline indicates is almost certainly a planet and 0.0 representing a TCE that the pipeline indicates is almost certainly a false positive. The disposition score for KOI 102.01 is 1.000, and that for KOI 102.02 is 0.4910, indicating that it is actually unclear whether it is correctly dispositioned as a false positive. 102.02 has a MES (multiple event statistic, see Eq. 5.12) of 22.4, compared to a MES for 102.01 of 446. KOI 102.02 is a TCE which I examined as part of my work with the *Kepler* False Positive Working Group, to confirm whether it is a false positive or not, and which I found to be a very interesting example of the problems which PLATO may face due to contamination by nearby stars.

[3] https://keplergo.arc.nasa.gov/ProposalPreparationApertures.shtml
[4] https://keplergo.arc.nasa.gov/DataAnalysisProducts.shtml

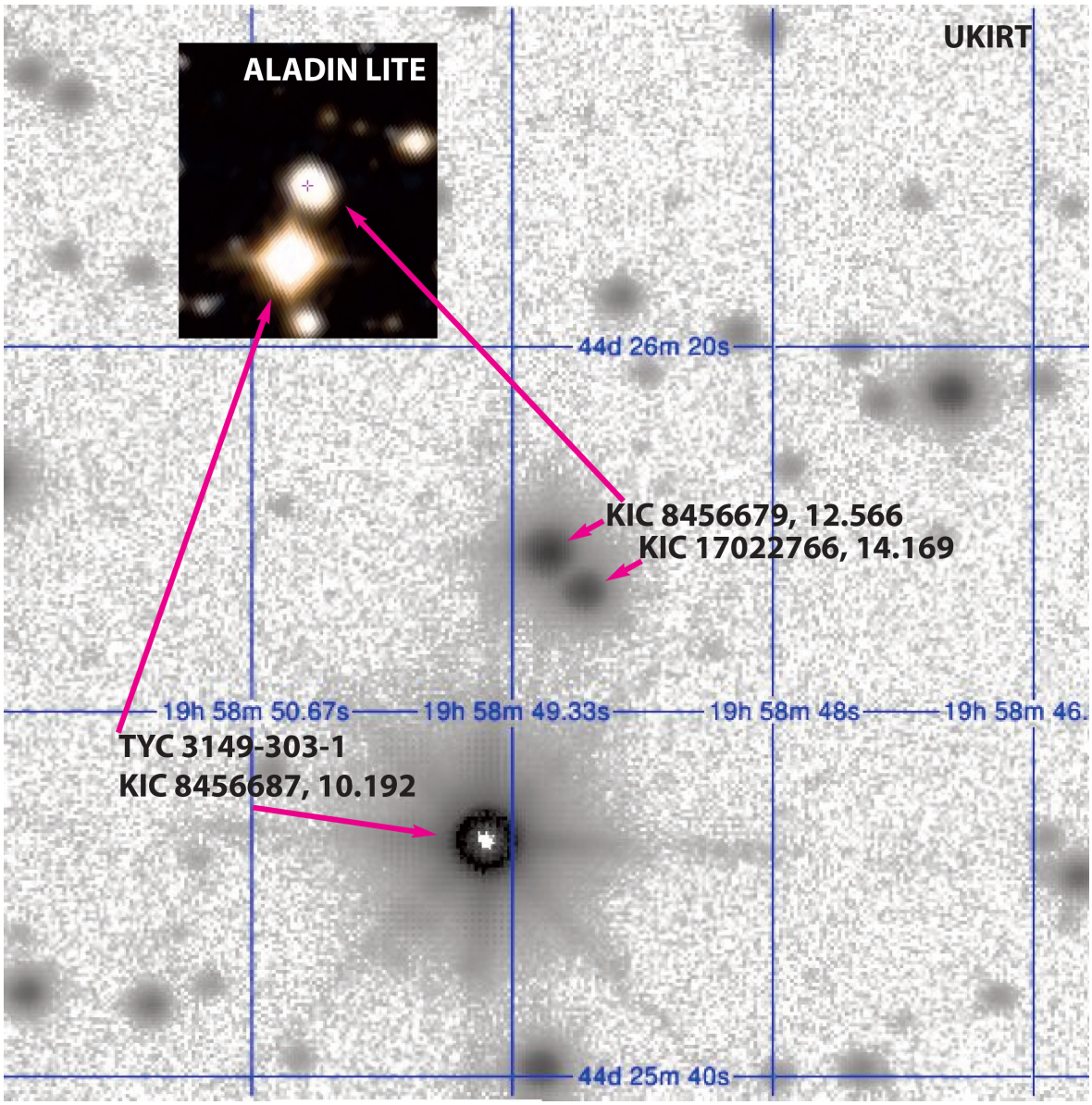


Fig. 5.6 KOI 102 (KIC 8456679) (Kp = 12.566): an example of a TCE whose analysis is affected by contamination from a bright nearby star. Main image: UKIRT in J band. Inset: Aladin Lite image (DSS). KOI 102 would have originally been believed to be one star: the later UKIRT image revealed the nearby star, KIC 17022766 (Kp = 14.169). Analysis of the two TCEs associated with KOI 102 is complicated by contamination from KIC 8456687 (Kp = 10.192): see text for further details. KIC 8456687 is 16.07" from KIC 8456679 and KIC 17022766 is 2.84" from KIC 8456679.

Analysis of this system is complicated by two factors, both of which are relevant for PLATO. First, the UKIRT image on the NASA ExoFOP (Exoplanet Follow-up Observing Program) website shows clearly a nearby star, 2.84" away (separation from ExoFOP), which, from its KIC number, was unlikely to have been separately observable before the UKIRT J-band image was taken. Evidence supporting this interpretation is the Aladin Lite image from Simbad, which shows the two stars, KIC 8456679 and KIC 17022766, blended as one. Gaia DR2 resolves two separate stars. For PLATO, these two stars will almost certainly fall on the same pixel. The second complication is the presence of TYC 3149-303-1 (also known as KIC 8456687), a Kp = 10.192 star (magnitude from MAST) 16.07" from KOI 102 (separation from Simbad), which in the *Kepler* data does not appear within the aperture around KOI 102, but which does appear to be contaminating pixels in or near the aperture. The aperture in *Kepler* is similar in purpose to the imagette in PLATO. While PLATO is unlikely to saturate at these magnitudes, KIC 8456687 will almost certainly be on a neighbouring PLATO pixel to KIC 8456679. From the data in the PLATO Definition Study Report and the Instrument Noise Budget, it appears likely that the PSF surfaces of these three stars will overlap. The relative positions of these stars are illustrated in Fig. 5.6, and illustrative difference plots (Q7) drawn from the publicly available TCERT report[5] are shown in Fig. 5.7 (aligned by column and row) and Fig. 5.8 (swung so aligned as in Fig. 5.6).

KIC 8456679 (KOI 102) would fall in the PLATO P5 population as it has a V-band magnitude of 11.95 (Simbad), while spectroscopic measurements give it an effective temperature of 5751^{+103}_{-114} K and a $\log(g)$ of $4.391^{+0.080}_{-0.150}$ (ExoFOP). Its distance from Earth is given as 368.89 ± 58.73 pc (ExoFOP), ≈ 400 pc (Gaia DR2). Centroid analysis in the *Kepler* pipeline indicates that the first TCE, with a period of 1.735 days, is probably on the target. The observing notes for this star on ExoFOP indicate that the star has narrow single lines, indicating that it is unlikely to be a short period binary. KIC 17022766 may be bound to KIC 8456679 in a wide orbit, or may be unbound: Gaia DR2 indicates a distance of ≈ 390 pc for KIC 17022766, and the error bars on the parallax of these two stars overlap in Gaia DR2. Both would fall on the same pixel as observed by PLATO: hence, it would be essential to take blending from KIC 17022766 into account when analysing any transit-like signals associated with KIC 8456679 in PLATO observations. There are no good *Kepler* difference images for the second, shallower TCE, with a period of 4.068 days. The offset is estimated at 2.034 ± 0.396 arcsec, approximately equidistant from KIC 8456679 and KIC 17022766, in a "triangular" shape, as shown in Fig. 5.7 and Fig. 5.8 (the triangle indicating the difference centroid is not over the x indicating target position for KOI 102.02, while the triangle is over

[5]https://exoplanetarchive.ipac.caltech.edu/data/KeplerData/008/008456/008456679/tcert/kplr008456679_q1_q17_dr25_obs_tcert.pdf

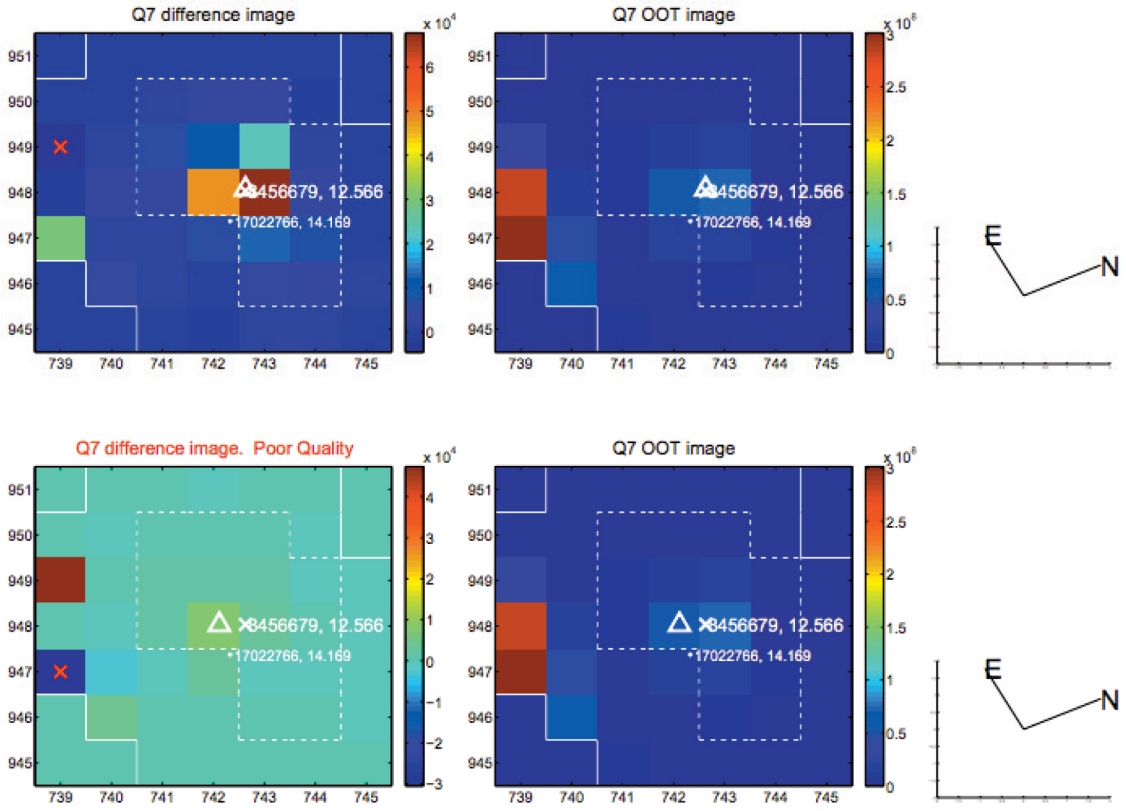


Fig. 5.7 KOI 102 (KIC 8456679) (Kp = 12.566) difference plots, Q7. Above, KOI 102.01, and below, KOI 102.02, aligned by column and row number. White x: KIC target position; +: OOT centroid; triangle: difference centroid, red x: large negative pixel value. Data from the publicly available TCERT report, https://exoplanetarchive.ipac.caltech.edu/data/KeplerData/008/008456/008456679/tcert/kplr008456679_q1_q17_dr25_obs_tcert.pdf

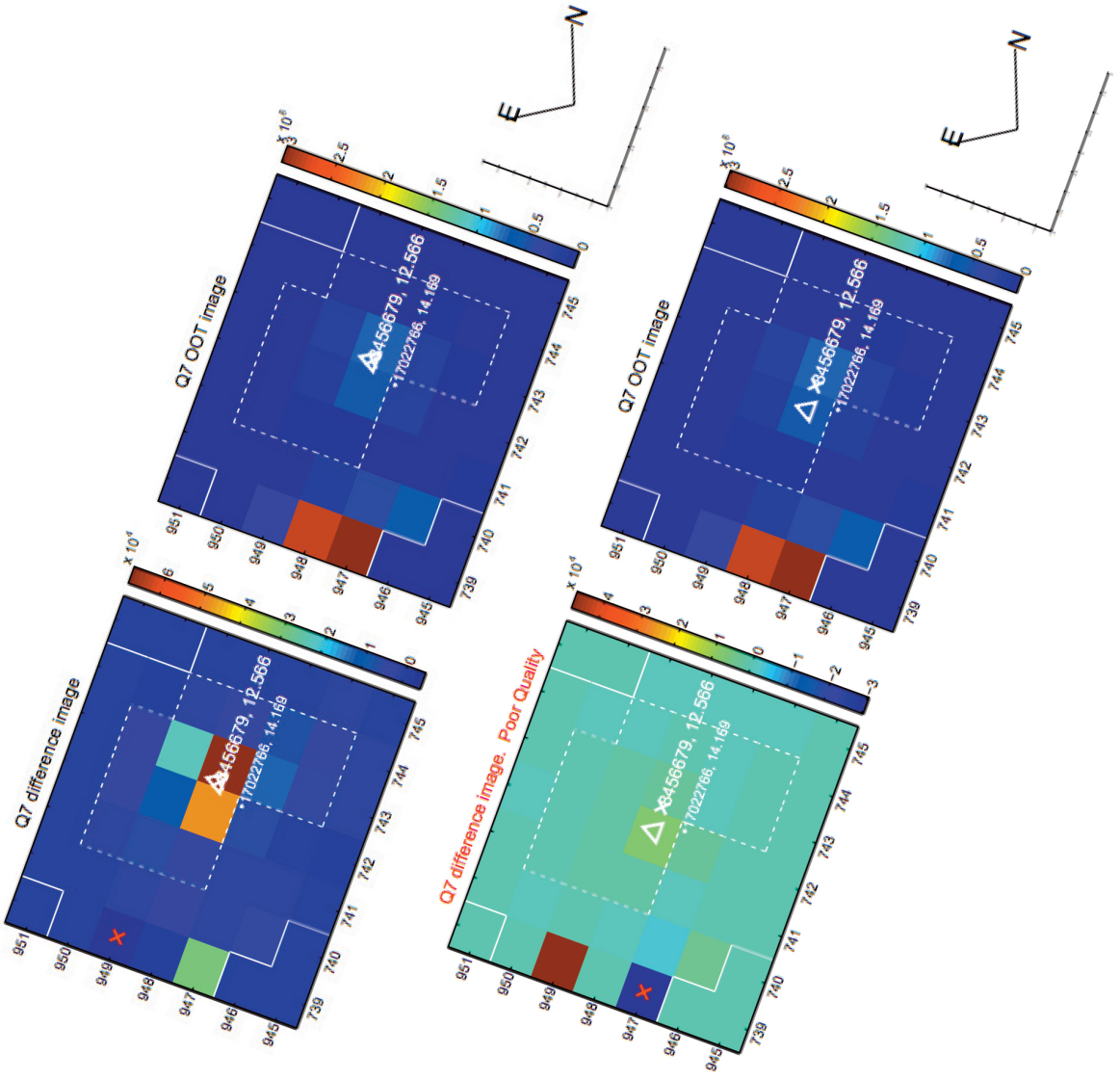


Fig. 5.8 As Fig. 5.7, swung so that the stars are aligned as in Fig. 5.6. Symbols as in Fig. 5.7

the x for KOI 102.01). The lack of good difference images makes it difficult to say if it falls on the target star, KIC 8456679, or not.

Dr Steve Bryson, Kepler Science Office Support Scientist, analysed a pixel level time series for this star, to see if this yielded additional information. Dr Bryson reported that this was inconclusive: due to the bright star, KIC 8456687, it was not possible to reliably compute centroids for 102.02, meaning that it was also not possible to determine where 102.02 was on the sky. A false positive determination due to a centroid offset could not be confirmed, but at the same time it could not be ruled out: all that could be reliably determined was that 102.02 was not centred on the bright star, KIC 8456687 (Bryson, priv. comm, 5 March 2018).

KIC 8456687 has a V-band magnitude of 10.47 (Simbad) but, as indicated in Fig. 5.6 where the pixels in the centre of the star have saturated, it is much brighter in the JHK bands (8.361, 7.787, 7.651) than it is in the B (11.76) and V (10.47) bands (Simbad). The distance in Gaia DR1 is $\approx$ 564 pc, and in Gaia DR2 is $\approx$ 640 pc. It is therefore probably further from Earth than the target star, KIC 8456679 (KOI 102), but brighter in both the Kp and V bands, with maximum flux in the infrared. This evidence all suggests that it is an evolved star, and therefore unlikely to be a PLATO target. However, flux from this star can be expected to contaminate any observations of KIC 8456679, in a neighbouring pixel, by PLATO.

Therefore, in addition to the issues addressed in this work concerning dilution by blending both of planet transits and binary eclipses, PLATO's large pixel size is likely to have a significant effect on target selection, as targets that stand proud within a pixel without bright stars in neighbouring pixels are likely to be favoured.

5.3 Summary: Chapter 5

With reference to the PLATO Instrument Noise Budget, PLATO's signal to noise characteristics have been reproduced in a simple model. With code provided by Catala (priv. comm), PLATO precision as a function of temperature and magnitude and of number of cameras can be applied to each synthetic system. As a consequence, the detectability of a given synthetic planetary transit or synthetic binary eclipse can be determined.

Issues arising from PLATO's large plate scale have also been examined, with a case study from *Kepler*, KOI 102.02, used to illustrate the issues that PLATO will face both from blending within the pixel and from contamination by nearby bright stars.

In the next chapter, I describe the process by which the intrinsic exoplanet distribution is estimated, and estimate the number of transiting exoplanets PLATO will observe in the two Long Look fields, in a one hour integration. This is an essential step in determining the effect of contamination on the PLATO planet haul.

Chapter 6

Intrinsic Exoplanet Distribution

While this project is mainly concerned with estimating numbers of blended and unblended eclipsing binaries, the relevance of this information in a transiting exoplanet survey depends on the relationship between the number of binaries and number of planets. Taking advantage of the synthetic Galaxy generated as part of this project (Chapter 4), estimated intrinsic planet distributions were obtained from *Kepler* observations. This allowed synthetic planetary populations to be created, and then interpreted using PLATO's signal to noise characteristics (Chapter 5).

This chapter describes how estimated intrinsic exoplanet distributions were obtained by calibrating to the *Kepler* planet haul, determines what can and cannot be learned about the true exoplanet distribution from these results, and presents predicted PLATO planet hauls for two different fields, one centred on $l = 65^\circ$, $b = 30^\circ$ and the other centred on $l = 253^\circ$, $b = -30^\circ$: the proposed Long Look North (LLN) and Long Look South (LLS) fields from Rauer et al. (2014). The Northern field includes the *Kepler* main mission field within it.

6.1 Generating the intrinsic exoplanet distribution

6.1.1 Biases and methods

The catalogue of 2293 planets (May 2017) discovered by the *Kepler* main mission and listed in the NExSci cumulative table is biased in several ways. One factor applicable to all exoplanet surveys is the transit probability: a planet with less than a given critical angle of inclination will not pass between its host star and an Earth-based observer, so will not be detectable.

Another bias is introduced by the instrument precision and signal-to-noise characteristics, described on the *Kepler* Guest Observer page[1] and the *Kepler* Instrument Handbook[2] (Van

[1] https://keplergo.arc.nasa.gov/CalibrationSN.shtml
[2] https://archive.stsci.edu/kepler/manuals/KSCI-19033-001.pdf

Cleve and Caldwell, 2016) (Batalha et al. (2010b), Gilliland et al. (2015)). These biases will be different for PLATO (Chapter 5). The passband will be different as well (Chapter 4) and, as a result, it is not possible to simply take the planet haul from *Kepler* and apply it to PLATO: it is necessary to establish from *Kepler* observations what the underlying distribution of planets is likely to be and then apply that to the context of PLATO, with its predicted signal to noise characteristics (Chapter 5).

Several methods have been used by various groups to determine the intrinsic exoplanet distribution from observations. Howard et al. (2012) combined information on early *Kepler* planet candidates and spectroscopy with transit probabilities to derive an intrinsic planet distribution within 0.25 AU of the host star, a method known as the "inverse detection efficiency method" or IDEM. Traub (2016) represented the frequency of exoplanets by a smooth function of planet radius and period, while Hsu et al. (2018) utilised an approximate Bayesian computation (ABC) to respond to issues presented by planets in the low signal to noise regime and at the limits of detectability, such as long period/small radius planets.

The method used in this work is distinct in that it combines a high quality synthetic Galaxy with a Monte Carlo method to achieve a binned intrinsic distribution, building on ideas in Fressin et al. (2013). This is closer to ABC than IDEM. As Hsu et al. (2018) discussed, published work based on IDEM tends not to estimate the intrinsic exoplanet population in regions in the *log P-log R* plane where planets have yet to be detected. This is something that has been attempted in this work. Hsu et al. (2018) reported the maximum priors which, using Bayesian techniques, led to no planets being simulated in the relevant bins. My constraints are unlikely to be as strongly defined as those used in Hsu et al. (2018), as my model is much simpler and is intended as an estimate only, but the same principal, of finding a PDF which reports planets in populated bins and does not report planets in unpopulated bins when compared to the observed *Kepler* population, is applied here.

The starting point is the *Kepler* photometric precision, designated in the *Kepler* Instrument Handbook (Van Cleve and Caldwell, 2016) as P, along with the geometric transit probability. P is derived in the *Kepler* Instrument Handbook as:

$$P = \sqrt{P_0^2 + \sigma^2} \tag{6.1}$$

where σ represents stellar variability, and

$$P_0 = p\frac{\sqrt{f_{kp}t_{int} + n_a N_R^2}}{\sqrt{n_{frames}} f_{kp} t_{int}} \tag{6.2}$$

where p is an empirical constant with a value between 1 and 2, t_{int} is the per-frame integration time = 6 seconds, n_a is the number of pixels in the photometric aperture, N_R is

Table 6.1 *Kepler* sensitivity as reported on the *Kepler* Guest Observer Page (single long cadence) and the *Kepler* Instrument Handbook (benchmark estimate). The benchmark set in the *Kepler* Data Processing Handbook is 20 ppm for a 12th magnitude star at a rms photometric precision appropriate for a hypothetical 6.5 hr integration.

Magnitude *Kp*	single long cadence (ppm)	benchmark estimate (ppm)
12.0	80.5	22.7
12.5	107	28.9
13.0	147	38.1
13.5	206	52.1
14.0	296	73.7
14.5	436	107.3
15.0	656	160.2
15.5	1003	243.6
16.0	1552	375.7

read noise = 120 e^{-1}, n_{frames} is the total number of frames per observation and f_{kp} is the expected photocurrent of the source,

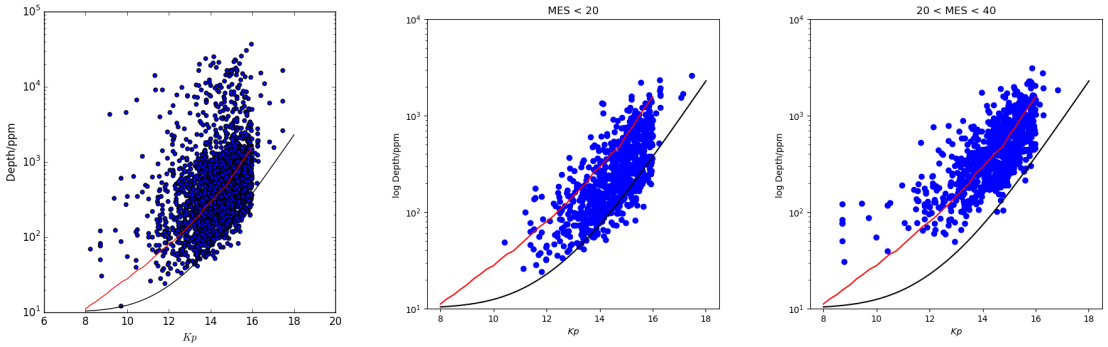
$$f_{kp} = 10^{-0.4(Kp-12)} f_{12} \tag{6.3}$$

where f_{12} is the benchmark photoelectron current at the focal plane for a G2 V star with $Kp = 12$. f_{12} was set pre-flight at 1.74 x 10^5 e^- s^{-1} with the option of updating in the light of operational experience (*Kepler* Guest Observer Page) and at 2.00 x 10^5 e^- s^-1 in the sample calculation in the *Kepler* Instrument Handbook.

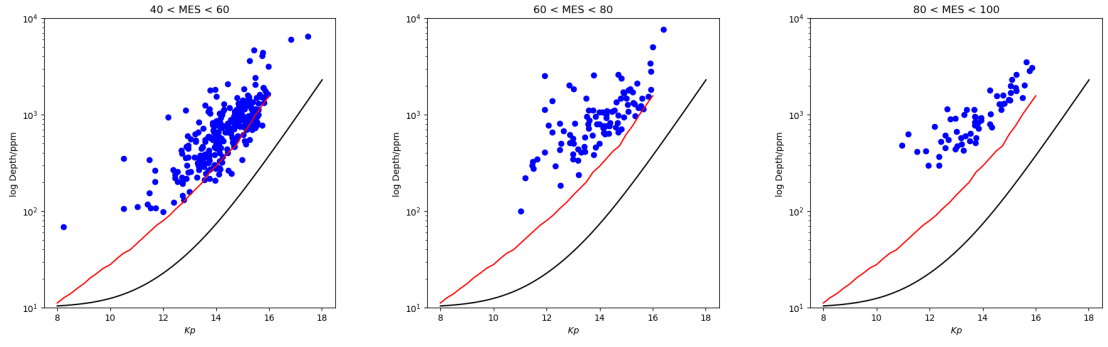
The estimate on the *Kepler* Guest Observer page for number of frames is 270 per 30 minute observation (one long cadence) and nine frames per one-minute observation (one short cadence). This allows for CCD readout time of 0.66 seconds. The *Kepler* Instrument Handbook includes a worked example for the benchmark estimate of the rms photometric precision at 6.5 hr, where the n_{frames} is 3575 and n_a is 20.

The empirical value, p, was set initially at 1.2, according to the *Kepler* Guest Observer Page, to be updated if required in the light of operational experience. The *Kepler* Instrument Handbook states that ≈ 1.2 was arrived at through modelling and lab photometry experiments. This empirical value accounts for all sources of noise not covered by other parts of the equation. Saturation begins in pixels containing stars with $Kp \leq 11$.

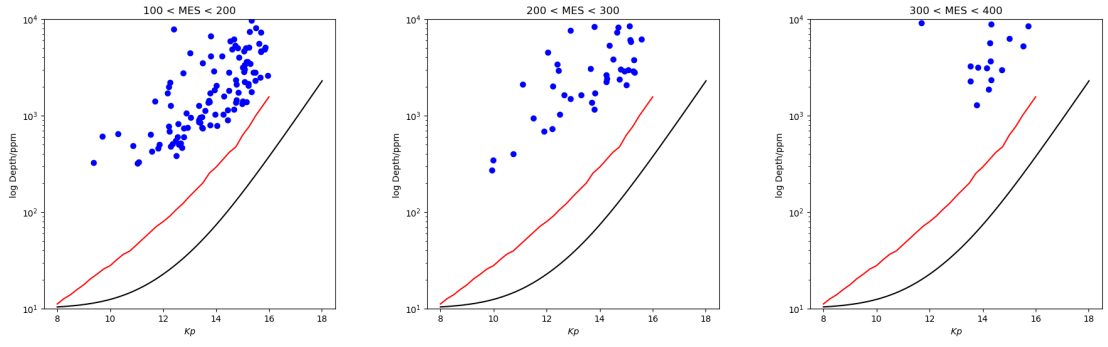
The *Kepler* Guest Observer page also demonstrates the match between pre-mission estimates and calibrated data obtained once the mission had began which shows that, for a 30 minute integration, the estimated precision was approximately correct.



(a) All confirmed *Kepler* planets. (b) 7.1 < MES < 20 (823). (c) 20 < MES < 40 (797).



(d) 40 < MES < 60 (265). (e) 60 < MES < 80 (100). (f) 80 < MES < 100 (61).



(g) 100 < MES < 200 (109). (h) 200 < MES < 300 (44). (i) 300 < MES < 400 (18).

Fig. 6.1 Confirmed planets, DR25, transit depth against magnitude. Red line: *Kepler* rms photometric precision of a single long cadence, black line: benchmark estimate. Panels (b) to (i) separate the data in panel (a) into bins by the recorded maximum multiple event statistic (MES) with the number of planets given in brackets. In addition, 72 planets have a maximum MES higher than 400. See Eq. 5.12 for the definition of the MES.

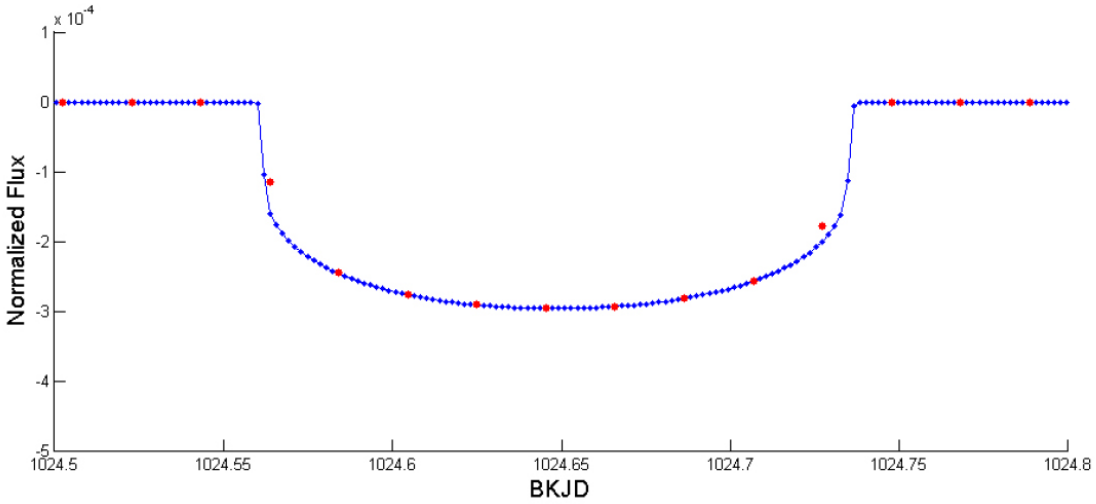


Fig. 6.2 Sample normalised *Kepler* light curve, Fig. 12.5 from the *Kepler* Data Processing Handbook (Li et al., 2017). Red points indicate the approximate mean of each long cadence interval, blue points indicate normalised flux at 1/11 of a long cadence ($\approx$ 2.67 min/point).

Fig. 6.1 shows all confirmed planets, *Kepler* DR25 (May 2017), transit depth against apparent magnitude, and compares these with the precision for a single long integration (panel a), using data from the *Kepler* Guest Observer Page (red line), and the 6.5 hr benchmark estimate, as quoted in the sample calculation in the *Kepler* Instrument Handbook (black line). Following the sample calculation, σ is set in this example at 10 ppm, although it is noted in the Handbook that 20 ppm may be more typical.

The process of identifying *Kepler* threshold crossing events (TCEs) is described in Seader et al. (2015). TCEs required a multiple event statistic, or MES (Eq. 5.12) of 7.1 to be considered for *Kepler* Object of Interest (KOI) status. Only KOIs were considered to be potential planets. From NExSci (accessed 4 November 2018), the lowest maximum MES of any confirmed planet is 7.127 (*Kepler* 445 d), and the lowest maximum single event statistic of any confirmed planet is 2.695 (*Kepler* 106 b, MES = 11.553). In terms of transit SNR (Eq.5.1), the single event statistic is equivalent to $N = 1$. This emphasises the point that every confirmed planet in Fig. 6.1 must have had at least one threshold crossing event where signal exceeded noise in a single long cadence integration. Panels (b) to (i) in Fig. 6.1 consider the MES of confirmed planets in depth-magnitude diagrams, relating this to the rms photometric precision of a single long cadence of average intrinsic noise and of the benchmark estimate. The lower envelope in panels (a), (b) and (c) appears to be approximated by the benchmark estimate.

Detection of signals by *Kepler* is based on long cadence data (Seader et al., 2015), and the *Kepler* Data Processing Handbook[3] demonstrates how exoplanet transits are extracted from the phase folded data, using priors. Fig. 6.2, Fig. 12.5 in the Handbook, provides a useful illustration of how this is achieved (Li et al., 2017). Each of the blue points in Fig. 6.2 represents the interpolated normalised flux at 1/11 of a long cadence, $\approx$ 2.67 mins, with the red points (the observed data) marking the 6th (blue) point in each long cadence. This example is *Kepler*-11 b, the sixth TCE of KIC 6541920. The priors reported by Li et al. (2017) in producing this curve are P = 10.30405 days, R_P/R_* = 0.0155697, a/R_* = 18.7471 and b = 0.1. With a transit depth of 316.1 $\pm$6 ppm and a host star of Kp = 13.709, *Kepler*-11 b is well within the single long cadence sensitivity (Table 6.1).

As demonstrated in Fig. 6.1 and in Table 6.1, when compared with PLATO sensitivity as described in Chapter 5, the *Kepler* benchmark estimate is more sensitive than the PLATO requirement in a one hour integration with 24 cameras in the centre of the field (34 ppm at V = 11 for stars for which asteroseismology data will be obtained). The *Kepler* Data Processing Handbook states that the *Kepler* requirement was for a rms photometric precision of 20 ppm for a 12th magnitude star in a benchmark 6.5 hr period. This is on the basis that if the Earth were to be observed transiting the Sun (12 hour transit), the transit depth would be 84 ppm.

Kepler observed the same field throughout its main mission, rolling once every 90 days to keep the Sun on the solar panels and the radiator cooling the focal plane pointed to deep space (*Kepler* Instrument Handbook). As a result of its four year observation programme, it includes data on planets with orbital periods of more than 365 days as well as planets with orbital periods of less than one day.

Fressin et al. (2013) studied the relationship between the intrinsic exoplanet population and that observed by *Kepler*, using data from Q1-Q6. In order to investigate the effect of blended planets on the false positive population, a single planet drawn from the then existing list of KOIs was inserted around a simulated background star matched to a star in the *Kepler* Input Catalogue. At the same time, to estimate the foreground population, a random planet was assigned to each *Kepler* target observed in Q1 to Q6 in five bins by planet radius and 11 bins by orbital period (see Table 6.2 and Table 6.3). The authors compared their simulated population of transiting exoplanets with the contemporary list of KOIs, minus their estimate of false positives, and adjusted their initial assumptions to match. It was recognised in Fressin et al. (2013) that the list of known KOIs was not unbiased for three reasons: because of transit probability, because it would contain an unknown number of false positives and because it would not account for the dependence of planet type on the nature of the host.

This study adopts similar principles to those used by Fressin et al. (2013), while taking advantage of the greater numbers of confirmed *Kepler* planets available today, to derive

[3]https://archive.stsci.edu/kepler/manuals/KSCI-19081-002-KDPH.pdf

Table 6.2 Radius bins used in Fressin et al. (2013), in $R_\oplus$

Earths	Super-Earths	Small Neptunes	Large Neptunes	Giant Planets
0.8-1.25	1.25-2.0	2.0-4.0	4.0-6.0	6.0-22.0

Table 6.3 Period bins used in Fressin et al. (2013), in days

0.8-2.0	2.0-3.4	3.4-5.9	5.9-10.0	10-17	17-29	29-50	50-85	85-145	145-245	245-418

estimated intrinsic exoplanet probability density functions (PDFs) based on both the *Kepler* single long cadence and the *Kepler* benchmark estimate, as approximate guides for planet detectability, for use with the PLATO simulations without having to include candidates that have not yet been confirmed as either genuine planets or false positives. It should be noted, however, that if all these unvalidated candidates are genuine planets, then the number of presently confirmed planets may represent only 51% of the true population of exoplanets observed and detected by *Kepler*.

6.1.2 Distribution of confirmed *Kepler* planets

The observed population of confirmed *Kepler* planets as described in the NExSci cumulative table, updated May 2017 and accessed May 2017 (DR25), was binned in two dimensions, by period and by planet radius, to obtain a target observed distribution in log space. Bins are 0.2 dex in width and cover the range $-0.4 < log\ P/\text{days} < 3.2$, $-0.4 < log\ R/R_\oplus < 1.6$. As described below, the normalised simulated output was checked against the normalised observed output at regular intervals. Because of the inherent statistical noise, and because some of the bins contain very few planets, variance between the simulated and observed population was expected. A total of 69 of the bins have no planets in them in DR25 (henceforward referred to as 'unpopulated bins'), either because no planets would be detectable due to the geometric transit probability (smaller planets at longer orbital periods, for example) or because planets would in theory be detectable at this radius/period, but none have yet been found. In the latter case, these bin counts may be zero due to effects of small number statistics (small planets at short periods), or there may be an astrophysical reason (see later references to the sub-Jovian desert, for example). In unpopulated bins, the objective was to reproduce this with no simulated planets, when using the *Kepler* precision.

Eq. 5.1 demonstrates that transit SNR improves as the number of events increases. Using data from Table 6.1, it is possible to estimate the maximum orbital period for which a transit SNR of 7.1 could be achieved for a planet of a given radius around a given host, as a function of magnitude and period. Table 6.4 does this for a Sun-like star and four sample planets:

Table 6.4 Theoretical detectability of a *Kepler* planet with a transit SNR $\geq$ 7.1 as a function of period and magnitude, in four years of observations. Data as presented relates to the period bins used in this work, and is max log P_{orb} where transit SNR $\geq$ 7.1. Noise is rms photometric precision, single long cadence. Four planets are considered against a Sun-like host: Jupiter (transit depth 10100 ppm), Neptune (transit depth 1250 ppm), a 2 $R_\oplus$ planet (transit depth 336 ppm) and Earth (transit depth 84 ppm). This is assessed using Eq. 5.1. Red indicates that the calculated rms photometric precision exceeds the transit depth for this model.

Magnitude *Kp*	Noise/ppm	Jupiter log *P*/days	Neptune log *P*/days	2 $R_\oplus$ planet log *P*/days	Earth log *P*/days
12.0	80.5	$>$ 3.2	$>$ 3.2	2.8	1.6
12.5	107	$>$ 3.2	$>$ 3.2	2.6	1.4
13.0	147	$>$ 3.2	$>$ 3.2	2.2	1.0
13.5	206	$>$ 3.2	3.2	2.0	0.8
14.0	296	$>$ 3.2	2.8	1.6	0.4
14.5	436	$>$ 3.2	2.4	1.4	0.2
15.0	656	$>$ 3.2	2.2	1.0	-0.2
15.5	1003	$>$ 3.2	1.8	0.6	$<$ -0.4
16.0	1552	$>$ 3.2	1.4	0.2	$<$ -0.4

Jupiter, Neptune, a 2 $R_\oplus$ planet and Earth. This table demonstrates that, even where the transit depth is less than the calculated rms photometric precision for a single long cadence, it is still possible to have a transit SNR > 7.1. All confirmed *Kepler* planets listed on NExSci have a MES $\geq$ 7.1. The sample that I have used in this work only includes confirmed planets listed on NExSci. Bins which, in period-radius log space, can only be populated by planets with a MES $<$ 7.1 will be unpopulated in the observed sample which I am trying to match.

Planets with periods $\geq$ 795 days cannot achieve the three detections necessary to be confirmed as a planet from *Kepler* data alone, given that *Kepler* obtained 1590 days of observations before the second reaction wheel failed. 530 days is one third of 1590 days, and 795 days is one half: this upper limit would allow for a detection to be made at the very beginning of the survey, at the very end and in the middle. On 11 April 2018 there are only four confirmed planets listed on NExSci with periods > 500 days: *Kepler*-1630 b (P = 510 days) (Morton et al., 2016); *Kepler*-548 b (P = 572 days) (Wang et al., 2015); *Kepler*-421 b (P = 705 days) (Kipping et al., 2014) and *Kepler*-167 e (P = 1071 days) (Kipping et al., 2016).

A feature noted in Fulton et al. (2017) was the presence of an "evaporation valley" separating rocky from gaseous planets, revealed by the California *Kepler* Survey (of which Fulton et al. (2017) is the third paper), which is becoming known as the "Fulton gap". The presence of such a feature is revealed in the DR25 data when binned by radius with a bin width of 0.05 dex at 0.25 < *log* $R/R_\oplus$ < 0.30, although it may be necessary to know what you

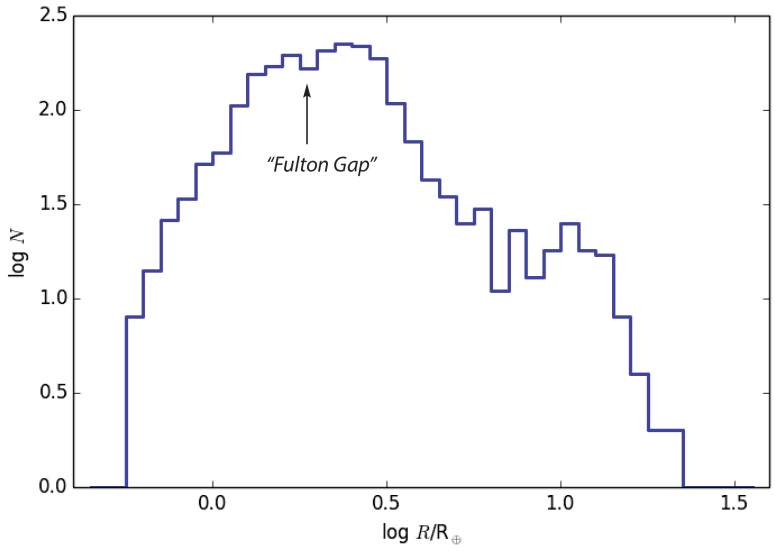


Fig. 6.3 *Kepler* observed planet distribution, DR25, in radius bins of 0.05 $R_\oplus$. The location of the "Fulton gap" or "evaporation valley" can be seen in the slight dip at $0.25 < log\ R/R_\oplus < 0.3$. The argument is that this gap separates terrestrial planets from those with a H/He atmosphere (Fulton et al., 2017).

are looking for to find it in the full data set (Fig. 6.3). Reference to Fulton et al. (2017) and their Fig. 2 shows that the gap was revealed after a series of cuts had been made to the original sample of candidates, removing false positives, and including only planets with brighter hosts ($Kp < 14.2$), low or intermediate impact parameters ($b < 0.7$) and with $P < 100$ days, where the host is a main sequence star with an effective temperature of $4700 < T_{eff}/\mathrm{K} < 6500$. While the "evaporation valley" or "Fulton gap" is of interest when considering questions of planetary architecture and planetary evolution, in a study such as this, where the aim is to reproduce the bulk planet population in a simple model, it is unlikely to be reproduceable.

The NExSci cumulative table does not include all *Kepler* planets listed on other websites, such as exoplanet.eu[4]. The most notable absence in NExSci is circumbinary planets in a P-type orbit, that is where the planet orbits both stars (often a close binary) rather than just one star. The values given for orbital period and, more commonly, radius are not necessarily the same in different catalogues, but the divergence between NExSci and exoplanet.eu has reduced considerably since the release of DR25, and where there is a discrepancy, the value in one catalogue is often within the error bars of the other catalogue. The decision was made to use data from NExSci because it is self-consistently derived from the original *Kepler* data, and therefore forms one complete data set.

[4]http://http://exoplanet.eu/catalog/

The synthetic *Kepler* field described in Farmer et al. (2013) was used to derive the intrinsic exoplanet distribution. This field had been carefully matched to the original *Kepler* target list in two senses: on sky field and target priorities. To match the on sky field, careful consideration was made of the area that each CCD would cover. This was an instantaneous snapshot, which did not take account of the seasonal rolls, in the same way that `GUMS` is an instantaneous snapshot of the field that would be observed by Gaia (Section 1.6.2). As with the *Kepler* target list, a shortlist of stars was selected from the full field, representing the candidates most likely to have detectable planet signals. Factors taken into consideration included stellar classification and signal to noise. Binaries were treated as point sources.

The NExSci cumulative table includes 2296 planets in 1632 systems (2 April 2018) (three had been added since the data used in this project was downloaded). 1201 (73.6%) of the systems contain one confirmed planet, 276 (16.9%) contain two confirmed planets, 99 (6.1%) contain three confirmed planets, 38 (2.3%) contain four confirmed planets, 16 (1.0%) contain five confirmed planets and two (0.1%) contain six confirmed planets. Therefore, 71.1% of the total number of confirmed planets carry the designation 'b', ie the first planet discovered in that system. Other planets may have been detected in these systems, but have yet to be confirmed and hence are listed only as candidates. Those remaining as candidates include those for which it was not possible to rule out a false positive scenario in Morton et al. (2016), where the term 'false positive' includes genuine planets orbiting stars other than the target. In the light of the figures concerning multiplanet systems, it was felt reasonable to follow Fressin et al. (2013) and to seed one synthetic planet onto each single star in the synthetic field.

As mentioned above, Fressin et al. (2013) discuss three concerns with regards to using the then contemporary KOI database to seed their synthetic population:

Geometric transit probability: As discussed in Chapter 3 in the context of eclipsing binaries, there is a critical angle of inclination below which an eclipse, or a transit, is unobservable for a given observer. The critical angle, i_c, is

$$i_c = \arccos\left(\frac{R_1 + R_2}{a}\right) \tag{6.4}$$

where R_1 and R_2 are the radii of the two bodies, both stars in the case of a binary and one star and one planet ($R_* + R_P$) in this context, and a is the semi major axis. Any transiting exoplanet survey favours larger planets closer to smaller stars, ideally in systems edge-on to the observer.

False positive content of the sample field: When Fressin et al. (2013) were working, the list of confirmed *Kepler* planets was significantly smaller than it is at present. To obtain a sufficiently large sample, Fressin et al. (2013) needed to work with the list of candidates, which would include genuine planets, false positives and false alarms. To account for this,

Fressin et al. (2013) also obtained a predicted false positive rate, which was applied to their synthetic output to reduce it to a more realistic size.

When work on the exoplanet distribution first begun in this study, in August 2015, the number of confirmed planets exceeded a thousand, but was still considered to be insufficient to achieve a true distribution without also considering unvalidated candidates, and therefore also needing to consider how many of these candidates were actually false positives and false alarms. The publication of Morton et al. (2016), as its title suggests, was a game changer: "False Positive Probabilities for all *Kepler* Objects of Interest: 1284 Newly Validated Planets and 428 Likely False Positives". With the validation of these 1284 planets, DR24 now contained 2290 confirmed planets, a sufficient number to ensure that only validated planets would be included in our primary calibrations and reducing the risk of counting a false positive as a planet. A few planets have been added since the publication of Morton et al. (2016) (on 12 March 2018 the total stood at 2296), but it does seem that Morton et al. (2016) provides the last significant word on which of the remaining candidates can be shown to be genuine planets. As previously noted, the updated stellar parameters in DR25 changed our understanding of the *nature* of these planets, but not on whether a given object is a planet, a false positive, a false alarm, or a candidate which is probably a planet but which cannot be confirmed as such at our present level of knowledge.

Dependence of planet type on the nature of the host: Fressin et al. (2013) began by assuming no dependence of planet type on the nature of the host. Those authors found that the assumption was valid, with the exception of their Giant Planet class (6-22 $R_{\oplus}$), where they state that they simulated an excess around M-class stars. Following Fressin et al. (2013) we also started with no dependence. Our results, given later in this chapter and in Appendix C, indicate that very few giant planets are simulated in PLATO priority class P4, M class stars with $V \leq 16$, but that in the most general class, $V < 26$, there is considerable dilution by blending of giant planets at $16 < V < 26$. This is a topic recommended for future study, in work more focused on the true exoplanet occurrence rate.

Fressin et al. (2013) also considered eccentricity, both for planets and for binaries. With planets, Fressin et al. (2013) assigned a random eccentricity and longitude of periastron using eccentricities drawn from a Rayleigh distribution with a mean value $e = 0.175$, for stars of all spectral types. The motivation for this was that eccentricity will alter the geometric transit probability and the duration of a transit.

We have simplified our model by assuming planet orbits are largely circularised. Incorporating eccentricity, and therefore considering the vectors along the line of sight to the observer and the line connecting the star and planet at periastron, is something that can be explored in future work more focused on the exoplanet distribution and occurrence rate.

A further simplification was to use only the single star synthetic population, leaving simulation of planets around binaries for future work. The *Kepler* pipeline uses the assumption that *Kepler* planet hosts are single stars (Ciardi et al. (2015) and references therein). In the synthetic *Kepler* field described in Farmer et al. (2013), $\approx 45\%$ of the synthetic systems are single stars (Section 4.2).

6.1.3 Calibration of the intrinsic exoplanet distribution

The objective in this study was to match a normalised intrinsic exoplanet distribution, $\mu_{\rm int}$, with the normalised *Kepler* observed distribution, $\mu_{\rm K}$. *Kepler* data was downloaded from NExSci, binned in the *log P-log R* plane as described in the previous subsection, and normalised, generating $\mu_{\rm K}$. Prior to normalisation, each bin in the observed population had a suitably small floor value (10^{-7}) added to it, to avoid having a zero probability of synthesising a planet in bins that were unpopulated in the *Kepler* data, when this may simply have been due to noise.

To avoid imposing assumptions on $\mu_{\rm int}$, the initial normalised distribution was flat: so

$$\mu_{\rm int,1}(log\ P,\ log\ R) = \text{constant} \tag{6.5}$$

One planet drawn randomly from $\mu_{\rm int,1}$ was seeded on each of the 94,355 single stars in the field described in Farmer et al. (2013), having also been assigned a random inclination. Undetectable synthetic planets, whether through transit probability or *Kepler* precision, or both, were discarded, as were any planets which would have been inside the star. The radius and period of detectable synthetic planets was saved. The seeding was repeated 200 times and summed over all 200 runs to reduce random noise.

The planets from the 200 seeding runs were binned in the *log P-log R* plane and the bins were normalised, generating $\mu_{\rm obs,1}$. Before normalisation, a suitably small floor value (10^{-7}) was added to the synthetic population, to avoid dividing by zero when the next PDF was generated:

$$\mu_{\rm int,2} = \mu_{\rm int,1}\frac{\mu_{\rm K}}{\mu_{\rm obs,1}} \tag{6.6}$$

So if the simulation was producing too many planets in a given bin, for example, the corresponding value in $\mu_{int,j+1}$ was reduced. The process was then iterated: so

$$\mu_{int,j+1} = \mu_{int,j}\frac{\mu_K}{\mu_{obs,j}} \tag{6.7}$$

Using an initially flat distribution resulted in an initial excess of giant planets which, in later iterations, disappeared. A synthetic population which approximated the observed

Table 6.5 Summary of labels for the different intrinsic exoplanet PDFs. 'Zeta' and 'omega' were generated for use in exoplanet simulations: 'sigma', 'tau', 'chi', 'omicron' and 'psi' were generated to test various aspects of 'zeta' and 'omega'. See the text for further information.

Label	Long cadence or benchmark estimate	Content of matching field
'zeta'	Long cadence	Confirmed planets
'omega'	Benchmark estimate	Confirmed planets
'sigma'	Benchmark estimater	Confirmed planets and candidates
'tau'	Benchmark estimate	Confirmed planets and candidates excluding objects with estimated radii > $10^{1.4}$ $R_{\oplus}$
'chi'	Benchmark estimate	Confirmed planets, populated bins only
'omicron'	Benchmark estimate	As 'chi' but with a 10x finer resolution
'psi'	Benchmark estimate	Smoothed distribution in two dimension based on confirmed planet distribution

population was achieved more quickly in a test which started with a distribution reflecting μ_K and the geometric transit probability, but I was concerned that this may have imposed a prior "signature" on μ_{int} and therefore was more comfortable with a flat initial distribution.

The process was undertaken several times utilising different parameters. The different PDFs, the integration times, the populations used and their labels, are summarised in Table 6.5. Various statistical methods were employed to test the simulated populations, including root mean square and χ^2 (global and normalised by the number of degrees of freedom). The PDFs 'zeta' and 'omega' were used to simulate exoplanets in the PLATO simulations. The remaining PDFs were used to explore the accuracy of 'zeta' and 'omega'.

In the observed population, 2285 confirmed planets and 2120 unconfirmed candidates fell within the bins used in this work (2071 unconfirmed candidates in 'tau' and 2308 'planets' in the smoothed population, 'psi'). To calibrate the PDFs, each bin in the synthetic population was multiplied by the total observed population and divided by the total synthetic population. The difference between the observed population and the synthetic population could then be calculated on a bin by bin basis, allowing goodness of fit to be calculated, as summarised in Table 6.6.

The initial set of seeding runs, using the flat population, was a poor fit in all the simulations. Too many short period, large planets were produced. In each case, the PDF quickly settled so that, in terms of root mean square, after a few runs little improvement was seen. The goodness of fit is consistently better in 'omega' than 'zeta'. This is likely to be a reflection of the fact that fewer planets were considered detectable in 'zeta'. Goodness of fit was improved in 'zeta'

Table 6.6 Statistical analysis, exoplanet PDFs, based on 200 seeding runs. Column 1: PDF. Column 2: RMS, populated bins. Column 3: RMS, unpopulated bins. Column 4: overall RMS. Column 5: Number of planets simulated in 200 seeding runs. Column 6: Normalisation factor. Column 7: χ^2, populated bins only. Column 8: As column 7, normalised to the degrees of freedom.

PDF	Populated RMS	Unpopulated RMS	All RMS	No. of planets	Normalisation factor	Global χ^2	Normalised χ^2
'zeta'	1.56	0.03	1.22	31261	14.6	9.4	0.085
'omega'	0.45	0.03	0.35	304435	1.5	1.4	0.013
'sigma'	1.08	0.06	0.99	176014	5.0	6.4	0.042
'tau'	0.97	0.04	0.90	173866	5.0	6.0	0.043
'chi'	0.47	0.02	0.37	301536	1.5	1.1	0.010
'omicron'	0.50	0.0044	0.39	307044	1.5	1.2	0.011
'psi'	0.59	0.0017	0.49	249112	1.9	1.9	0.016

by combining 10 sets of 200 seeding runs, producing a single data set of 2,000 independent seeding runs. Similar improvements were noted when, in testing of the final PDF in 'omega', a total of 1,200 independent seeding runs were considered. This indicates that 200 seeding runs were insufficient to fully eliminate statistical noise. However, increasing the number of seeding runs beyond 200 had implications in terms of computing time.

With 'omega', shallow transits on faint stars were more likely to be considered detectable than with 'zeta'. This is responsible for the increased numbers of planets detected in equivalent runs using 'omega', and the lower normalisation factor required to match the observed *Kepler* population.

Many unconfirmed candidates represent objects with shallow transit depths around faint stars which are difficult, if not impossible, with our current level of technology to confirm. They do not appear to be false positives or false alarms, but a false positive or false alarm scenario may not be able to be ruled out. The PDF 'sigma' was obtained to understand the effect that including these unconfirmed candidates would have. It should also be noted that the list of unconfirmed candidates also includes a number of objects which are probably too large to be planets (> 25 $R_\oplus$), and therefore might be brown dwarfs or binaries with M-class secondaries not recognised by the *Kepler* pipeline as being stellar. Objects of this size are treated with caution in 'zeta', 'omega' and 'sigma', but are included in 'sigma' for completeness. The PDF 'tau' was derived in order to understand the effect that bodies with an estimated radius $1.4 < log\ R/R_\oplus < 1.6$ may have on the PDF: three such bodies are confirmed planets, and a further 49 are unconfirmed candidates. The normalisation factor is similar in both 'sigma' and 'tau' and higher than in 'omega'. More terrestrial planets are included in the observed distribution in 'sigma' and 'tau' than in 'omega', although the *Kepler* sensitivity used to derive these PDFs is the same. This indicates that the normalisation factor increases

when the number of terrestrial planets in the observed population is high for the *Kepler* sensitivity with which the synthetic population is being examined; this will contribute to the high normalisation factor in 'zeta'.

'Chi' was obtained to examine the effect of simulating planets in bins which are unpopulated in the observed population, where this may simply be due to noise. This in turn raised issues of machine precision and led to the derivation of 'omicron' at a 10x finer resolution, discussed in the following subsection. 'Omega', 'omicron' and 'chi' have the same normalisation factor. 'Psi' was obtained to examine whether peaks and troughs in the observed population made reproducing these features in the simulations more complicated. Again, the results are examined in the next subsection.

All the PDFs represent estimates derived from simplified models and should be considered with this in mind.

6.1.4 Verification

Having obtained the test PDFs, further verification of their ability to satisfactorily reproduce the observed population was an essential next step.

The original *Kepler* target list consisted of 150,000 systems with $Kp \leq 16$ (Batalha et al., 2010b). This was updated as the mission progressed, with some targets being added and some being removed, and Hsu et al. (2018) estimates that $\approx$ 192,000 stars were observed in total. The field used in the simulations, from Farmer et al. (2013), was simulated using the same selection criteria as the original target list. Most of the observed *Kepler* exoplanets have hosts with $Kp < 16$, but there are 21 confirmed planets orbiting 13 host stars with *Kp* 16-18 (see Table. 4.2 for the magnitude distribution of exoplanet host stars).

The combined effect of the simplifications described above, simulating one planet per system, only using the single star population, the variation in the number of systems observed and the presence of planets around stars with $Kp > 16$, suggest our simulations would be expected to produce about one third of the number of observed planets, if the assumption that the synthetic *Kepler* field accurately reflects the composition of the observed *Kepler* field holds. Therefore, it was anticipated that the normalisation factor would be ≈ 3, if the observed *Kepler* population was complete. From Table 6.6, the normalisation factor for 'zeta' is 14.6, for 'omega', 'chi' and 'omicron' it is 1.5 and for 'sigma' and 'tau' it is 5.0. From the composition of the populations and the sensitivities on which these PDFs were based, I conclude that 'omega', 'chi' and 'omicron' are based on an observed population that is incomplete, but that including unvalidated candidates, as was the case with 'sigma' and 'tau', means the proportion of terrestrial planets in the observed population rises, and as a result the normalisation factor rises beyond the expected range.

Work was also undertaken to understand why the χ^2 values failed to improve after the first few simulations. Two questions were asked, both concerned with the possibility of overfitting the data: what happened if the unpopulated bins were set to zero probability, and what effect smoothing the distribution would have.

When 'chi' was run, it was immediately noticeable that planets were being reported in bins for which the probability of simulating a planet had been set at zero. Examining the boundaries of bins where this was occurring, the most likely explanation was machine precision. The bin boundaries and intervals were set in log space with full 32-bit precision, but the planet radii and orbital periods were being saved to only 12 significant figures. Where the bin boundary was *log x*.2 and *log x*.8, it appeared that synthetic planets on the boundary were being reported one bin smaller by radius or period than that in which they had been simulated. To test this, 'omicron' was run with a resolution 10x finer than 'chi', and 10x fewer planets appeared in unpopulated bins, in proportion to the bin they were shifted from, than in 'chi'.

In the populated bins, the PDF was able to adjust to this movement, increasing the probability in bins that were 'losing' planets and decreasing the probability in bins that were 'gaining' planets, so there is very little change from one set of runs to the next.

In the unpopulated bins affected by this issue, a small population would be consistent with the bins being unpopulated in the observed population due to noise. As the bins from which planets were being 'transferred' only had one or two planets in the observed population, it would be expected that only 0.05-0.10 planets per bin would be recorded, once the synthetic population was normalised to the observed *Kepler* population. Probably only the high number of seeding runs used in the process of deriving the PDF (200) allowed this issue to come to light.

The observed population was smoothed when working with 'psi', so that in every column and every row there was only one maximum. The smoothing was based primarily on the confirmed planet population, but the distribution of the candidate population was also taken into account. This 'observed' population was artificial, but the intention was to reduce the amount of noise that the PDF needed to work with. Bins which were anomalously populated in 'chi' and 'omicron' were now populated as part of the smoothing process. Fluctuation in the rms and χ^2 remained.

However, when combining the results of multiple sets of seeding runs conducted with the same PDF, as was the case with the final PDF in 'zeta' and 'omega', it was noticeable that all the goodness of fit measures improved. For example, χ^2 normalised by the number of degrees of freedom improved from 0.013 to 0.010 in 'omega' (6 x 200 seeding runs) and from 0.085 to 0.019 in 'zeta' (10 x 200 seeding runs). This leads to the conclusion that 200 seeding runs were insufficient to eliminate random noise, and emphasises the point that

these intrinsic exoplanet distributions are estimates. Increasing the number of seeding runs also increased the computing resources required, and in this sense 200 seeding runs was a reasonable compromise.

Similar starting distributions tended to produce similar outcomes. 'Omega', 'chi' and 'omicron' are similar, as are 'sigma' and 'tau'. All PDFs were derived independently from a flat distribution, with 'omega', 'chi' and 'omicron' matched to one observed distribution (confirmed planets) and 'sigma' and 'tau' matched to another (confirmed planets plus candidates). As is the case with confirmed planets, all candidates listed on NExSci with a recorded MES have a MES ≥ 7.1.

The similarity in 'omega', 'chi' and 'omicron', and in 'sigma' and 'tau', indicate that the method of obtaining an estimate of the intrinsic exoplanet population is sound, with the caveat that estimates of the distribution in bins that are unpopulated in the observed population should be regarded as possible maxima, not as final representations.

6.2 Discussion

Several physical factors influence the PDFs 'zeta' and 'omega'.

1. Only single stars were used in these simulations. To include binaries would require calculating the minimum stable orbit around both stars (P-type), and calculating the maximum orbit in the case of a planet orbiting one component of the binary (S-type). This is a recommendation for future work. While exoplanet.eu reports both P-type and S-type planets in binary systems, a comparison with the NExSci cumulative table shows that in the latter case P-type planets are excluded, even where the planet has an official *Kepler* name. This is due to the difficulty in disentangling the shallower, longer period signal from a planet from the deeper, shorter period signal from the binary (J. Coughlin, priv. comm). Examples include *Kepler*-47 (AB) b and *Kepler*-47 (AB) c (Orosz et al., 2012), discovered, according to the paper, by a synergy between the teams working on planet detections and eclipsing binary detections. So while it is valid not to consider P-type planets when comparing to the NExSci cumulative table, we are in effect allocating the S-type planets to the single star population, without allowing for possible dilution by blending. This is in line with the *Kepler* pipeline, which treats all *Kepler* candidate host stars as single stars (Batalha et al. (2011), Burke et al. (2014), Mullally et al. (2015), Ciardi et al. (2015)).

Similarly, dilution by blending with unbound background stars may also lead to an underestimate of the radius of a planet.

2. In the paper describing the original target list selection process (Batalha et al., 2010a), 41% of the 150,000 priority systems in the original *Kepler* target list have $15 < Kp < 16$. Among the single stars in the synthetic *Kepler* target field selected by Farmer et al. (2013) to

match the original *Kepler* target field, and used to derive the PDFs 'zeta' and 'omega', only 30% of the stars fall into this category (32% for all systems). Both the original *Kepler* target list and the field in Farmer et al. (2013) were magnitude limited to $Kp < 16$.

3. The *Kepler* catalogue of confirmed planets is not a complete record of all detectable planets transiting observed stars. For the catalogue to be complete, the following assumptions would need to be made: all transiting planets with a suitable transit SNR that have been observed in the field have been detected by the pipeline; and all candidates not yet confirmed as planets are false positives. These assumptions are highly unlikely to be true.

Further consideration of these points follows.

6.2.1 Dilution by blending in binaries and with background stars

Where the flux of more than one star, whether bound or unbound with the target star, needs to be taken into consideration when analysing events, any transit will be diluted. Binaries with planets in an S-type orbit are likely to have a long binary period, and there is no guarantee that the orbits are coplanar; so a transit by a planet does not imply that there will be an observable stellar eclipse in such systems. At the time of writing (14 September 2017) there are 22 known circumbinary planets in S-type orbits in the *Kepler* catalogue, or $\approx 1\%$ of the total: see Appendix B for a list of these planets, with references. Known S-type planets are not significant in the total *Kepler* population.

However, this only applies to systems known to be binaries. Consider a binary with a planet or planets in an S-type orbit in which the angular separation of the stellar components is (much) smaller than the pixel scale, and for whatever reason, whether it is long orbital period, an orbit that is not coplanar with the orbit of the planet(s) around the star, or stars with very unequal luminosities, the system is not recognised as a binary but is believed to be a single star. In such a situation, the observed flux would be from both stars and the true stellar parameters would be different from what they appear to be.

136 very short period (≤ 1 day) *Kepler* Objects of Interest (KOIs) believed to be eclipsing binaries with significant secondary eclipses from the NExSci cumulative table (latest update: DR24) were compared by me with the relevant entries from Gaia DR1, soon after the release of Gaia DR1. In most cases where a match was found, Gaia had also recorded a single object: this is expected as one of the stated limitations of DR1 is an inability to separate close binaries. The value given for the *Kepler* magnitude was also very similar to the value given for the Gaia magnitude. The exception was KOI 3616, listed in the NExSci cumulative table with $Kp = 15.839$. Two objects were found at the relevant co-ordinates by Gaia, one with $G = 16.626$ and the other with $G = 16.110$.

Consider two stars in the same pixel, in a non-eclipsing binary orbit, one with G = 16.626 and the other with G = 16.110. Using

$$m_{tot} = -2.5log(10^{-m_1/2.5} + 10^{-m_2/2.5}) \tag{6.8}$$

a combined magnitude m_{tot} of G = 15.585 is obtained. Using

$$\delta F_{obs} = \frac{\delta F_{act}}{10^{\Delta m/2.5}}. \tag{6.9}$$

where δF_{obs} is the observed eclipse/transit depth, δF_{act} is the actual eclipse/transit depth and Δm is the magnitude difference between m_{tot} and m_1 or m_2 as is applicable in context, a planet in an S-type orbit around the brighter star would appear to have a transit depth 62% of the true value, and a planet in an S-type orbit around the fainter star would appear to have a transit depth 38% of the true value. Such a dilution would certainly reduce the apparent size of the planet, and may even render it undetectable. A determination of the planet radius would of course be made with reference to the derived stellar radius which, in the case of an unresolved binary, may be incorrect, especially if based on photometry alone. If spectroscopy of the host star is available, this would be more reliable.

For such a planet, there are therefore two good reasons why the derived planet radius may be incorrectly derived: dilution by blending and incorrect stellar parameters.

Dilution by blending clearly applies to unbound stars in the same pixel, as well as to bound systems.

The *Kepler* Eclipsing Binary Catalogue (KEBC)[5] lists 2910 eclipsing binaries found in the *Kepler* output (accessed 9 April 2018, last updated 27 April 2017). 150,000 systems were included on the original *Kepler* target list (Batalha et al., 2010a) and ≈ 192,000 were observed at some point during the main mission (Hsu et al., 2018). Raghavan et al. (2010) find that 56 ±2% of systems in a volume-limited sample are single stars, with the rest being binaries and other hierarchies. So *Kepler* should have observed ≈ 85,000 multiple stellar systems. Most will not be eclipsing: Raghavan et al. (2010) found only three eclipsing binaries in a volume limited sample of 454 systems, of which 33±2% were observed to be binaries, 8±1% were observed to be triples and 3±1% were observed to be quadruple systems or higher hierarchies. Dilution by blending in a multiple stellar system clearly affects triples and higher hierarchies as well as binaries.

Ciardi et al. (2015) examined the effects of undetected stellar companions in the *Kepler* mission, with particular reference to applying lessons learned to K2 and TESS. Ciardi et al. (2015) define a planetary correction factor, X_R and find that, if KOIs are assumed to be single, the planetary radii may be underestimated by a factor of $< X_R > \approx 1.5$, a factor reducing to

[5]http://keplerebs.villanova.edu

$< X_R > \approx 1.2$ if high resolution imaging and RV work reveal no observable stellar companions. Ciardi et al. (2015) found that the correction factor $< X_R >$ is dependent on stellar properties, ranging from $< X_R > \approx 1.6$ for A and F class stars to $< X_R > \approx 1.2$ for K and M class stars. Ciardi et al. (2015) refines these headline figures by deriving $< X_R >$ as a function of T_{eff}:

$$< X_R >= a_3 \left(T_{eff}\right)^3 + a_2 \left(T_{eff}\right)^2 + a_1 \left(T_{eff}\right) + a_0 \tag{6.10}$$

where $a_3 = -1.19118$ x 10^{-11}, $a_2 = 1.61749$ x 10^{-7}, $a_1 = -0.000560$ and $a_0 = 1.64668$ (Equation 8 from Ciardi et al. (2015)). Further refinements derive the uncertainty in $< X_R >$ from the uncertainty in T_{eff}. $< X_R >= 1.0$ is the solution appropriate for genuine single star systems.

Hirsch et al. (2017) reported on a campaign in which 176 close (< 2") stellar companions detected near 170 KOI stars were examined using high-resolution imaging. Systems were selected for this work on the basis that 176 out of the 204 planets or planet candidates had estimated planet radii of < 6 $R_\oplus$. Hirsch et al. (2017) found that 60-80% of the companions within 1" were gravitationally bound, and that >90% of companions within 0.5" were bound. Using the assumption that planets in a bound binary system were equally likely to orbit the primary or the secondary, Hirsch et al. (2017) reassessed the estimated radius for these planets and found that $< X_R >= 1.65$. Hirsch et al. (2017) conclude that, unless vetted by high resolution imaging or spectroscopy, nearly half of all *Kepler* planets may have radii underestimated by an average of 65%.

Furlan et al. (2017) found, from high resolution imaging, that 10% of *Kepler* target stars have a companion within 1" and 30% have a companion within 4". Furlan et al. (2017) estimated that, as a result, the number of *Kepler* planets with a radius < 2 $R_\oplus$ should decrease by $\approx$ 2-23%.

Ciardi et al. (2015), Hirsch et al. (2017) and Furlan et al. (2017) all agree that the *Kepler* catalogue, using the single star assumption, underestimates the radius of a proportion of planets due to blending with bound and unbound companions. Dilution by blending can therefore be expected to influence intrinsic exoplanet distributions extrapolated from that catalogue, with Neptune-like planets appearing to be terrestrial and terrestrial planets diluted by blending to the point where they are no longer observable, in a proportion of cases. The true number of Neptune-like planets detectable by *Kepler* is therefore likely to be underestimated and there is a case for saying that the true number of terrestrial planets detectable by *Kepler* is unknown.

Table 6.7 Comparison of magnitudes of stars in the original *Kepler* Input Catalogue (Batalha et al., 2010a) with the magnitudes of stars in the synthetic *Kepler* field from Farmer et al. (2013). [1](Batalha et al., 2010a), [2]Farmer et al. (2013).

Kp	KIC[1]		Single[2]		Binary[2]		Total[2]	
3-4					1	0.0009%	1	0.0005%
4-5					3	0.0026%	3	0.0014%
5-6			7	0.0074%	8	0.0070%	15	0.0072%
6-7	5	0.0033%	29	0.031%	14	0.012%	43	0.021%
7-8	51	0.034%	62	0.066%	70	0.061%	132	0.063%
8-9	252	0.17%	201	0.21%	237	0.21%	438	0.21%
9-10	673	0.45%	598	0.64%	679	0.59%	1277	0.61%
10-11	1773	1.2%	1655	1.8%	1857	1.5%	3411	1.6%
11-12	4600	3.1%	3378	3.6%	3996	3.5%	7374	3.5%
12-13	12063	8.0%	9565	10%	10672	9.3%	20237	9.7%
13-14	30451	20%	21394	23%	24580	22%	45974	22%
14-15	38991	26%	28801	31%	34816	30%	63617	31%
15-16	61141	41%	28665	30%	37552	33%	66217	32%
Total	150000	100%	94355	100%	114384	100%	208739	100%

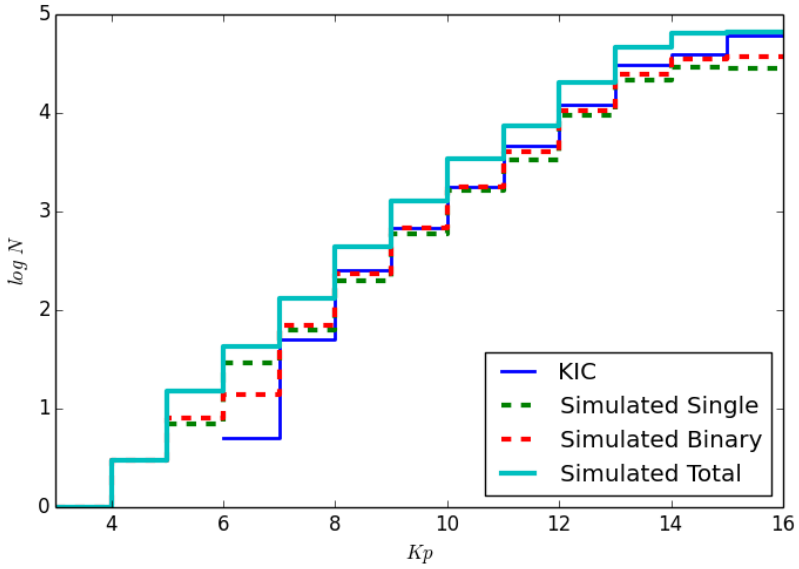


Fig. 6.4 Comparison of magnitudes of stars in the original *Kepler* Input Catalogue (Batalha et al., 2010a) (KIC) with the magnitudes of stars in the synthetic *Kepler* field from (Farmer et al., 2013) (simulated single, simulated binary and simulated total).

6.2.2 Magnitude distribution

In the field generated in Farmer et al. (2013), used to calibrate the intrinsic exoplanet distribution, stars with $15 < Kp < 16$ were under-represented compared to the original *Kepler* target list (Table 6.7 and Fig. 6.4), although both had the same magnitude limit ($Kp \leq 16$). This suggests that, in the range $14 < Kp < 16$, Farmer et al. (2013) selected more bright targets from a similar starting point than the original *Kepler* target list. Farmer et al. (2013) also include a few stars brighter than $Kp = 6$, the upper limit in the KIC, but these are unlikely to be statistically significant.

Table 6.7 and Fig. 6.4 compare the original *Kepler* Input Catalogue (KIC) (Batalha et al., 2010a) as a function of magnitude with the numbers of single stars, binary stars and all stars in the field from Farmer et al. (2013). Most *Kepler* planet hosts (34.2%) have magnitudes $15 < Kp < 16$ (Table 4.2, Fig. 4.6), but while such stars make up 41% of the original KIC, only 30% of single stars (32% of all stars) in the field from Farmer et al. (2013) have $15 < Kp < 16$.

Apparent magnitude is determined by three factors: intrinsic luminosity, distance and extinction. Distance from Earth is considered in Fig. 6.5, which shows that, using *Kepler* precision and a single long cadence observation (30 minute integration), there is, as expected, a clear relationship between the distance from Earth at which a planet can be observed and its radius: larger planets can be observed at a greater distance.

In a sample biased towards brighter stars, planets can be detected with a lower SNR in a transiting exoplanet survey (Fig. 6.1), and this is reflected in transiting exoplanet survey planning. Batalha et al. (2010a) indicated that stars brighter than $Kp = 14$ were favoured in the original *Kepler* target list, and both TESS and PLATO plan to target brighter stars in the solar neighbourhood, in part because such stars are more amenable to asteroseismology and spectroscopic follow up, but also for the simple reason that planets orbiting brighter host stars are generally easier to detect.

The difference in magnitude distribution at $14 < Kp < 16$ between the original *Kepler* input catalogue and the field from Farmer et al. (2013) would make it easier to detect transiting planets in the synthetic population, after allowing for the difference in numbers between the original *Kepler* Input Catalogue and the single stars in the field from Farmer et al. (2013).

6.2.3 *Kepler* completeness

In this work, we have used confirmed planets and planet candidates from *Kepler* DR25. How complete is this catalogue?

Christiansen et al. (2015) conducted an experiment to try to establish this. Synthetic transiting planet signals were injected into one year's worth of pixel data for $\approx$ 10,000 targets. Comparing the recovered signals, after normal processing, with expectations, allowed

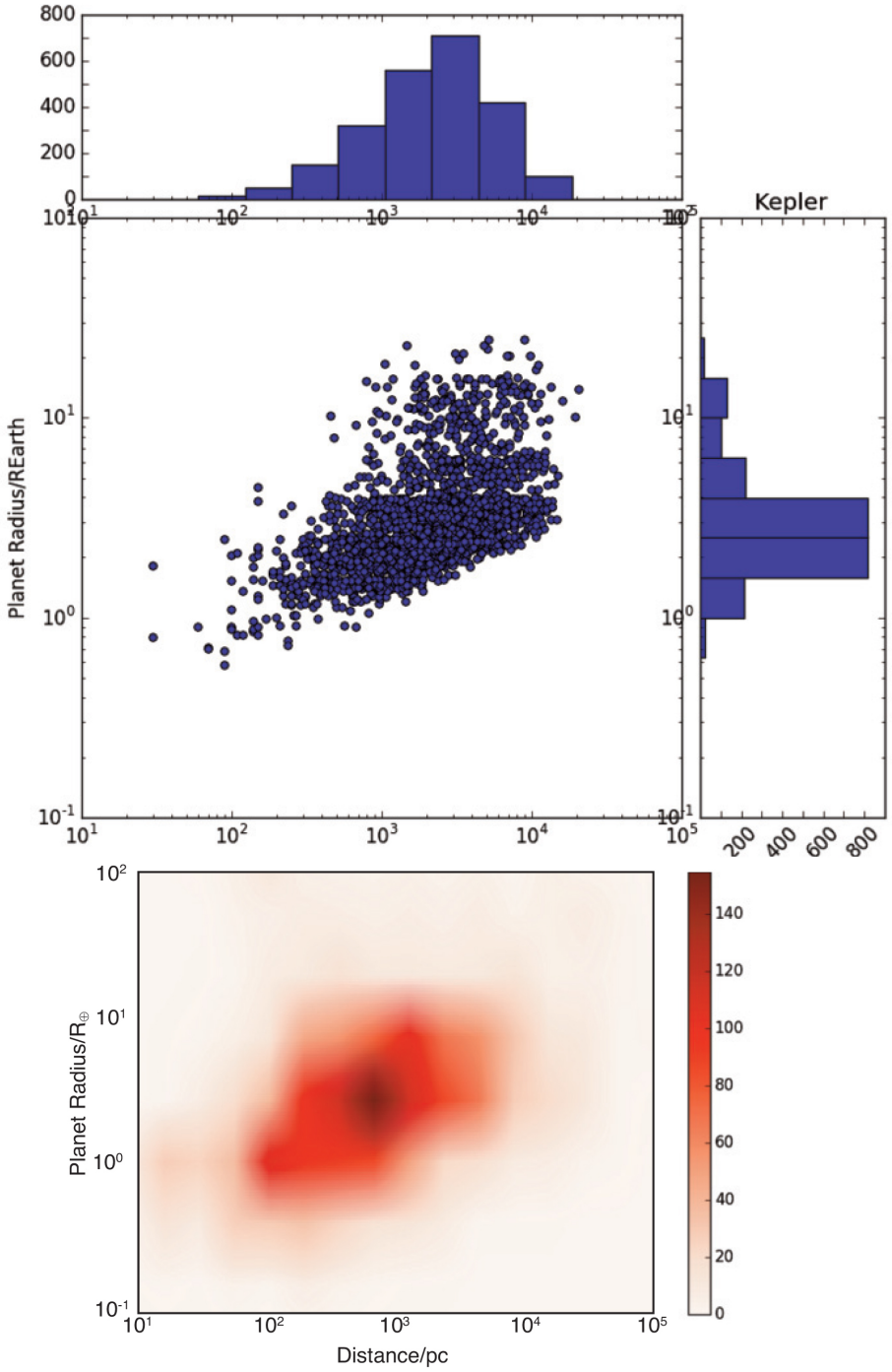


Fig. 6.5 A simulated planet population interpreted with *Kepler* precision (one long cadence), plotting planet radius against distance from Earth, in a 3° x 3° box around the equivalent of the centre of the *Kepler* main mission field, drawn from the field generated for the PLATO simulations. The lower panel is a binned contour plot. There is a clear relation between planet radius and the maximum distance at which a planet can be observed. Sample simulated with the 'zeta' PDF.

Table 6.8 Comparison of occurrence rates in Fig. 7, Christiansen et al. (2015), their Γ function over perfect sensitivity, by radius and period.

Radius $R_\oplus$	0.5-1.25 days	1.25-2.5 days	2.5-5 days	5-10 days	10-20 days	20-40 days	40-80 days	80-160 days	160-320 days
1.75-2	n/a	1.00	1.03	1.05	1.08	1.14	1.19	1.25	1.30
1.5-1.75	1.01	1.03	1.05	1.09	1.16	1.21	1.26	1.30	1.36
1.25-1.5	1.00	1.05	1.12	1.16	1.22	1.27	1.31	1.38	1.40
1-1.25	1.08	1.14	1.20	1.25	1.29	1.33	1.38	1.40	n/a

Table 6.9 Comparison of occurrence rates in Fig. 7, Christiansen et al. (2015), ramp function from Fressin et al. (2013) over perfect sensitivity, by radius and period.

Radius $R_\oplus$	0.5-1.25 days	1.25-2.5 days	2.5-5 days	5-10 days	10-20 days	20-40 days	40-80 days	80-160 days	160-320 days
1.75-2	n/a	1.06	1.10	1.16	1.24	1.44	1.58	1.79	1.79
1.5-1.75	1.04	1.06	1.17	1.27	1.46	1.67	1.83	2.00	2.27
1.25-1.5	1.05	1.15	1.35	1.46	1.68	1.87	2.04	2.38	2.54
1-1.25	1.24	1.42	1.60	1.78	1.86	2.13	2.35	2.55	n/a

the construction of a sensitivity curve of signal recovery as a function of signal to noise. Christiansen et al. (2015) stated that while the sensitivity from these experiments did not match the theoretical curve, the result was "not as pessimistic as some published estimates had stated". For example, the recovery rate of 7.1σ signals was 25.8%, compared to a theoretical 50% recovery rate. The results of the experiment are illustrated in their widely quoted Fig. 3.

Christiansen et al. (2015) found that the signal recoverability of the pipeline was well characterised by a Γ cumulative distribution function of the form

$$p = F(x|a,b) = \frac{1}{b^a \Gamma(a)} \int_0^x t^{a-1} e^{-5/b} \mathrm{d}t \tag{6.11}$$

where $a = 4.35$, $b = 1.05$ are the best-fit co-efficients for FGK stars, and $a = 4.77$, $b = 1.24$ are the best fit coefficients for other stars.

To illustrate the effect this may have on occurrence rates, Christiansen et al. (2015) computed toy model occurrence rates for FGK stars for planets $1 < log\ R/\mathrm{R}_\oplus < 2$ with periods $0.5 < P/\mathrm{days} < 320$ using the perfect signal recovery, the sensitivity described in Eq. 6.11 and the ramp function from Fressin et al. (2013), all as illustrated in their Fig. 3, and presented the results in their Fig. 7. Taking the ratios of the occurrence rates, as shown in Table 6.8 and Table 6.9, illustrates that smaller planets and planets at longer periods are most likely to be affected by a lack of completeness.

Table 6.10 Estimated planet radius, *Kepler* DR25, confirmed planets and unconfirmed candidates, DR25 (May 2017).

Radius/log $R/R_\oplus$	Confirmed	%	Candidates	%	Total	%
−0.4 to −0.2	10	0.44	66	3.1	76	1.7
−0.2 to 0.0	126	5.5	278	13	404	9.2
0.0 to 0.2	490	21	604	29	1094	25
0.2 to 0.4	797	35	538	25	1335	30
0.4 to 0.6	584	26	293	14	877	20
0.6 to 0.8	133	5.8	99	4.7	232	5.3
0.8 to 1.0	65	2.8	76	3.6	141	3.2
1.0 to 1.2	68	3.0	63	3.0	131	3.0
1.2 to 1.4	9	0.39	54	2.6	63	1.4
1.4 to 1.6	3	0.13	49	2.3	52	1.2

Fig. 6.1 shows that most confirmed *Kepler* planets have transit depths larger than the benchmark estimate described in the *Kepler* Instrument Handbook. However, confirmation comes as a result of follow up work, not directly through the *Kepler* pipeline. Some candidates are simply unsuitable for spectroscopic or photometric follow up, due to the faintness of the host star. Morton et al. (2016) validated a large number of planets by ruling out false positive scenarios: this leaves an even larger number of unvalidated planet candidates, for which a false positive scenario could neither be confirmed or discounted.

Table 6.10 demonstrates that unvalidated candidates listed on NExSci are more significant than confirmed planets in the three smallest radius bins used in this work, as well as the largest two bins. Confirmed planets are more significant at $0.2 < log\ R/R_\oplus < 0.8$ (Neptune-like planets).

Table 6.11 considers transit duration of candidates and confirmed planets. Transit duration tends to increase with orbital period, as shown in Fig. 6.6 and Fig. 6.7. There are a significant number of candidates with transit durations > 10 hours (10%), which suggests these candidates also have longer periods, although there are some outliers at shorter periods, which may be orbiting red giants. The longer period and longer transit duration is likely to make confirmation through observation more difficult.

The work by Christiansen et al. (2015) demonstrates that the theoretical precision is not achieved in injection-and-recovery tests. Therefore, it is likely that the true intrinsic exoplanet distribution lies between the PDF 'zeta', which required a higher minimum transit depth for a planet to be detectable at a given magnitude and orbital period, and the PDF 'omega', which is affected by completeness issues. 'Zeta' indicates higher intrinsic numbers of terrestrial planets, due to the higher detection threshold, while 'omega' increases the proportion of giant planets. Fig. 6.8 demonstrates this.

Table 6.11 Transit duration, *Kepler* DR25, confirmed planets and unconfirmed candidates (8 April 2018).

Transit duration/hr	Confirmed	%	Candidates	%	Total	%
< 0.5	2	0.09	10	0.45	12	0.26
0.5-1.0	14	0.61	76	3.4	90	2.0
1.0-1.5	100	4.4	181	8.1	281	6.2
1.5-2.0	200	8.7	220	9.8	420	9.2
2.0-2.5	252	11	233	10	485	11
2.5-3.0	281	12	226	10	507	11
3.0-3.5	267	12	210	9.4	477	11
3.5-4.0	192	8.4	171	7.6	363	8.0
4.0-4.5	177	7.7	120	5.3	297	6.5
4.5-5.0	161	7.0	109	4.9	270	5.9
5.0-5.5	112	4.9	93	4.1	205	4.5
5.5-6.0	91	4.0	64	2.9	155	3.4
6.0-6.5	85	3.7	62	2.8	147	3.2
6.5-7.0	72	3.1	57	2.5	129	2.8
7.0-7.5	64	2.8	34	1.5	98	2.2
7.5-8.0	33	1.4	36	1.6	69	1.5
8.0-8.5	30	1.3	37	1.6	67	1.5
8.5-9.0	32	1.4	34	1.5	66	1.5
9.0-9.5	15	0.65	22	0.98	37	0.81
9.5-10.0	16	0.70	20	0.89	38	0.84
> 10.0	101	4.4	229	10	330	7.3

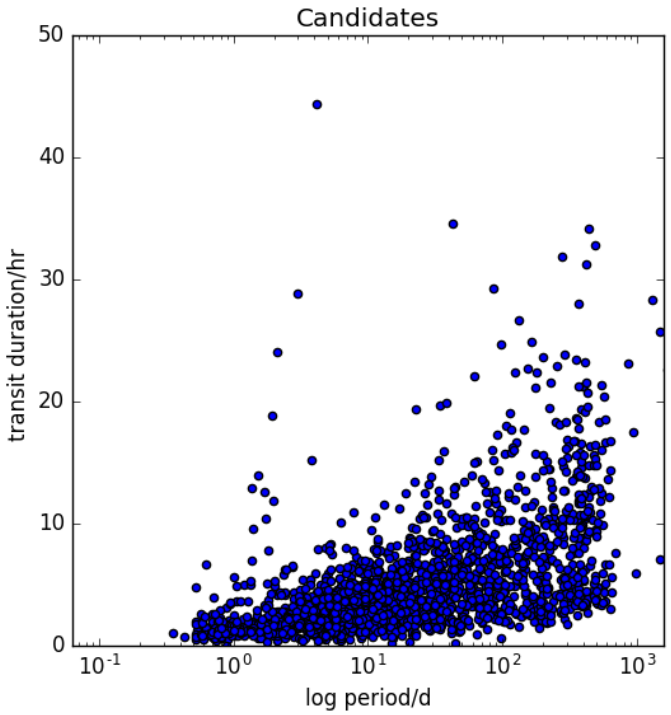


Fig. 6.6 Unvalidated candidates, *Kepler* DR25, transit duration against orbital period.

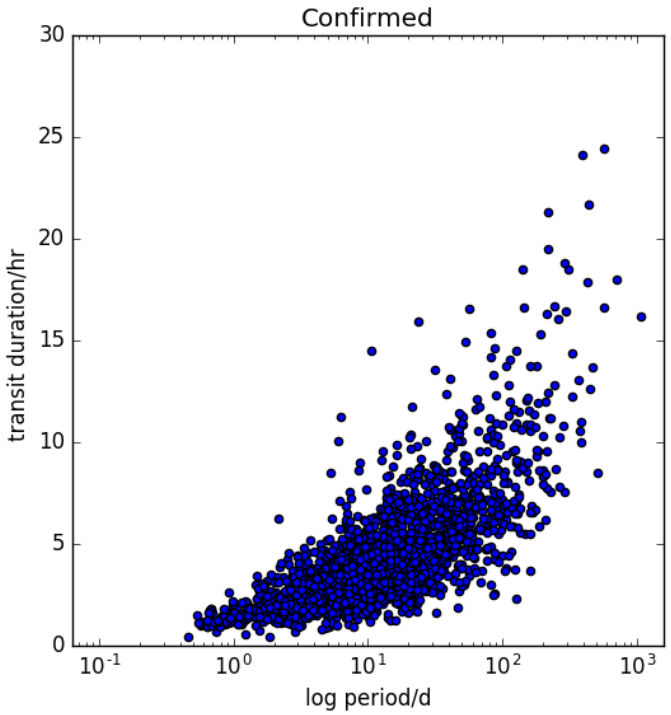
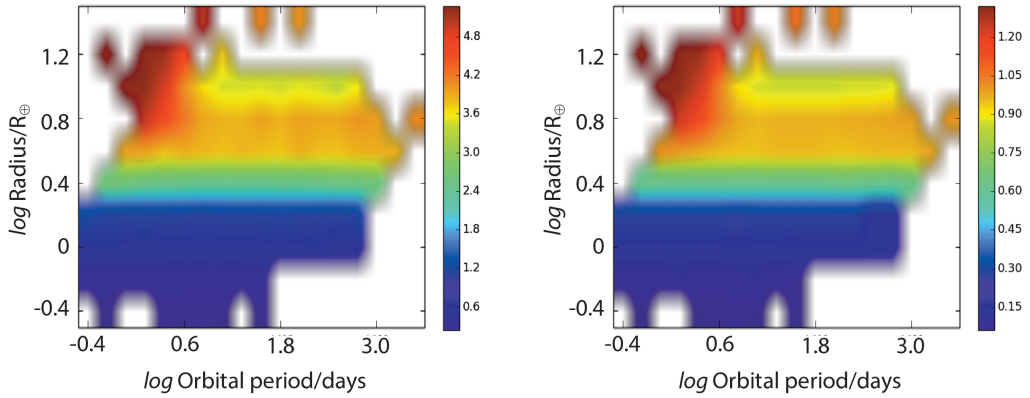


Fig. 6.7 Confirmed planets, *Kepler* DR25, transit duration against orbital period.



(a) PDF used in simulations: 'zeta' (b) Flat PDF

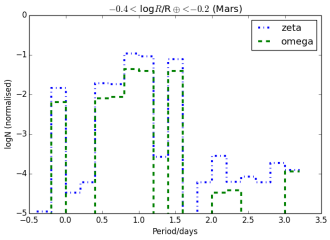
Fig. 6.8 Ratio of planets observable in the same synthetic population by *Kepler*, 30 min sensitivity to 6.5 hr sensitivity. Each simulation consisted of 2000 seeding runs. The 'zeta' PDF was used to derive the distribution analysed in panel (a), and a flat PDF was used to derive the distribution in panel (b). The synthetic populations were normalised to the *Kepler* confirmed planet population before the ratios were taken. Note the difference in scale on the colour bars in the two panels.

In effect, 'zeta' takes into account the completeness issues which 'omega' raises.

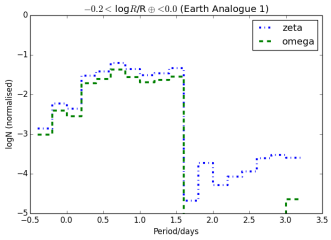
6.3 Analysis of the estimated intrinsic planet distributions

The PDFs 'zeta' and 'omega' are illustrated in Fig. 6.9 and Fig. 6.10, and are given in Table C.1, with values for 'omega' in black and 'zeta' in red in the table. Only 'zeta' and 'omega' are used in the PLATO exoplanet simulations, so only these two PDFs are discussed in detail in this section.

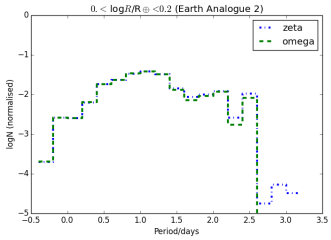
Fig. 6.9 demonstrates that in populated bins, 'zeta' and 'omega' have very similar features but are offset from one another. This is a natural consequence of the fact that both have been fitted to match the same observations. This demonstrates the applicability of my method, given that both PDFs were derived from an initially flat distribution. In the radius bin $0 < log\ R/R_{\oplus} < 0.2$ (panel (c)), 'zeta' and 'omega' are remarkably similar. At smaller radii, 'zeta' indicates a higher intrinsic distribution than 'omega', and at larger radii, 'omega' indicates the higher intrinsic distribution. Sudden dips are explained by the presence of unpopulated bins in the observed population. All panels have been fixed to have the same *y* axis, to enable easier comparison between panels. Panel (a) represents the smallest planets included in this work, with several bins that may have been unpopulated due to statistical noise, and therefore the distribution is not as smooth as the other panels. The sharp drop offs noted in each panel at longer periods, especially at smaller radii, are a consequence of the fact that I was only trying



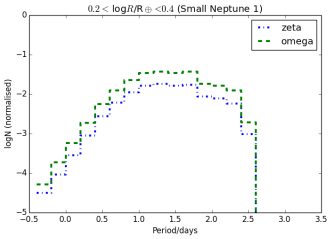
(a) $-0.4 < logR/\mathrm{R}_{\oplus} < -0.2$



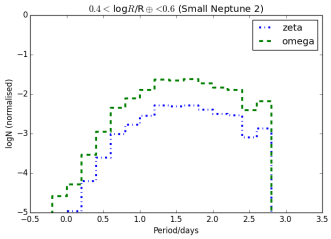
(b) $-0.2 < logR/\mathrm{R}_{\oplus} < 0.$



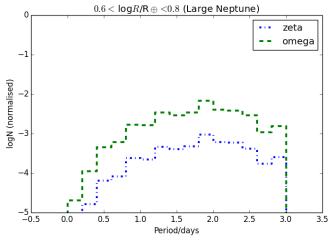
(c) $0. < logR/\mathrm{R}_{\oplus} < 0.2$



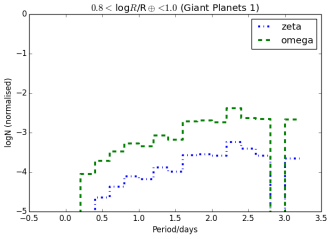
(d) $0.2 < logR/\mathrm{R}_{\oplus} < 0.4$



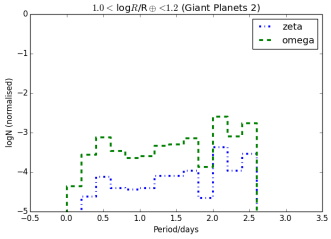
(e) $0.4 < logR/\mathrm{R}_{\oplus} < 0.6$



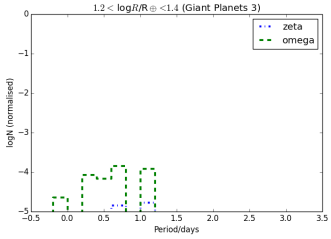
(f) $0.6 < logR/\mathrm{R}_{\oplus} < 0.8$



(g) $0.8 < logR/\mathrm{R}_{\oplus} < 1.$



(h) $1. < logR/\mathrm{R}_{\oplus} < 1.2$



(i) $1.2 < logR/\mathrm{R}_{\oplus} < 1.4$

Fig. 6.9 The PDFs 'zeta' and 'omega' by radius bin, normalised to 1, with the *y* axis presented in log scale. The intrinsic distribution in the radius bin $0 < logR/\mathrm{R}_{\oplus} < 0.2$ is remarkably similar in both PDFs in bins which are populated in the observed distribution.

to match planets with a MES $\geq$ 7.1: so the intrinsic distribution is likely to be underestimated in such bins.

Fig. 6.10 uses contour mapping to compare the populated bins in the observed distribution of confirmed planets with equivalent bins in 'zeta' and 'omega', and Fig. 6.11 highlights features in the observed distribution which are described in more detail in Section 6.3.1: the sub Jovian desert, the three-day pile-up, the maximum radius as a function of insolation and the region below the minimum detection threshold.

As expected, small planets are intrinsically more common than larger planets in both 'omega' and 'zeta' (Fig. 6.9). The results for each planet class, based on radius, are discussed in more detail later in this section.

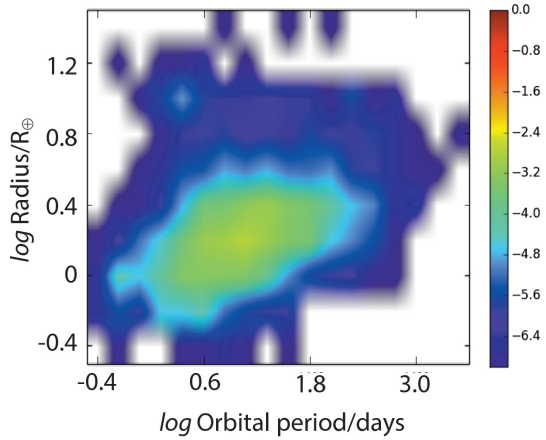
Fig. 6.12 maps the confirmed planets (panel a) and synthetic populations derived using 'zeta' (panel b) and 'omega' (panel c), as well as the absolute difference between the observed and synthetic populations in log space (panels d and e). (a), (b) and (c) appear very similar. The difference between the observed and simulated populations is lowest in the region that is most densely populated by detectable planets.

To enable comparison of my results with work utilising the commonly used bins from Fressin et al. (2013), populations are reported in Section 6.4 in both logarithmic bins and bins from Fressin et al. (2013). To see how these differing systems of bins compare, Fig. 6.13 illustrates the observed distribution in DR25, with logarithmic bins as used in this work in an underlying grid, in black, and the period and radius bins from Fressin et al. (2013) overprinted in red. While the planet classes from Fressin et al. (2013) are based on physical properties of the planets, they are not evenly distributed in either linear or logarithmic space.

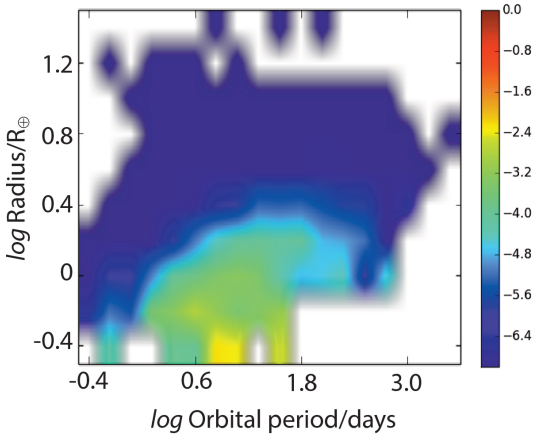
The aim in finding the intrinsic exoplanet distribution was to be able to apply it to the PLATO simulations. Therefore, as further validation of 'zeta' and 'omega', I examined the distribution of synthetic planets in two simulations drawn from a PLATO synthetic field, consistent with the on-sky region observed with *Kepler*, including all main sequence single stars with $Kp < 26$. One simulation was generated with the 'zeta' PDF and interpreted with *Kepler* sensitivity for a 30 minute integration, the other was generated with the 'omega' PDF and interpreted with the *Kepler* sensitivity for the benchmark estimate described in the *Kepler* Instrument Handbook.

Blending of planet hosts with nearby stars was taken into account in both simulations by considering the contribution of all sources within a given pixel, whether resolvable or not, and whether single stars or binary systems. Observations of planets around faint stars will naturally be more affected by blending than observations of planets around bright stars.

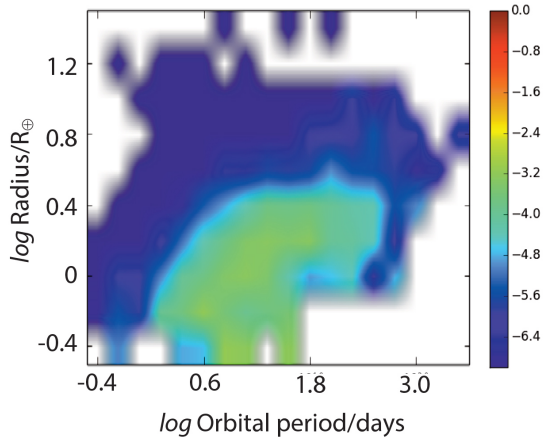
Table 6.12 presents the results of this analysis, also illustrated in Fig. 6.14 ('zeta'). In both simulations, more planets are considered detectable by *Kepler* before blending is taken into account and after accounting for the difference in area than are recorded as confirmed



(a) Observed planets (normalised), DR25.



(b) 'zeta' PDF, normalised, populated bins only.



(c) 'omega' PDF, normalised, populated bins only.

Fig. 6.10 Contour maps of the observed confirmed exoplanet distribution and the derived intrinsic distributions, over planet radius and orbital period. (a) Confirmed *Kepler* planets (normalised), (b) the intrinsic distribution in the 'zeta' PDF and (c) the intrinsic distribution in the 'omega' PDF. Scale bar is logarithmic.

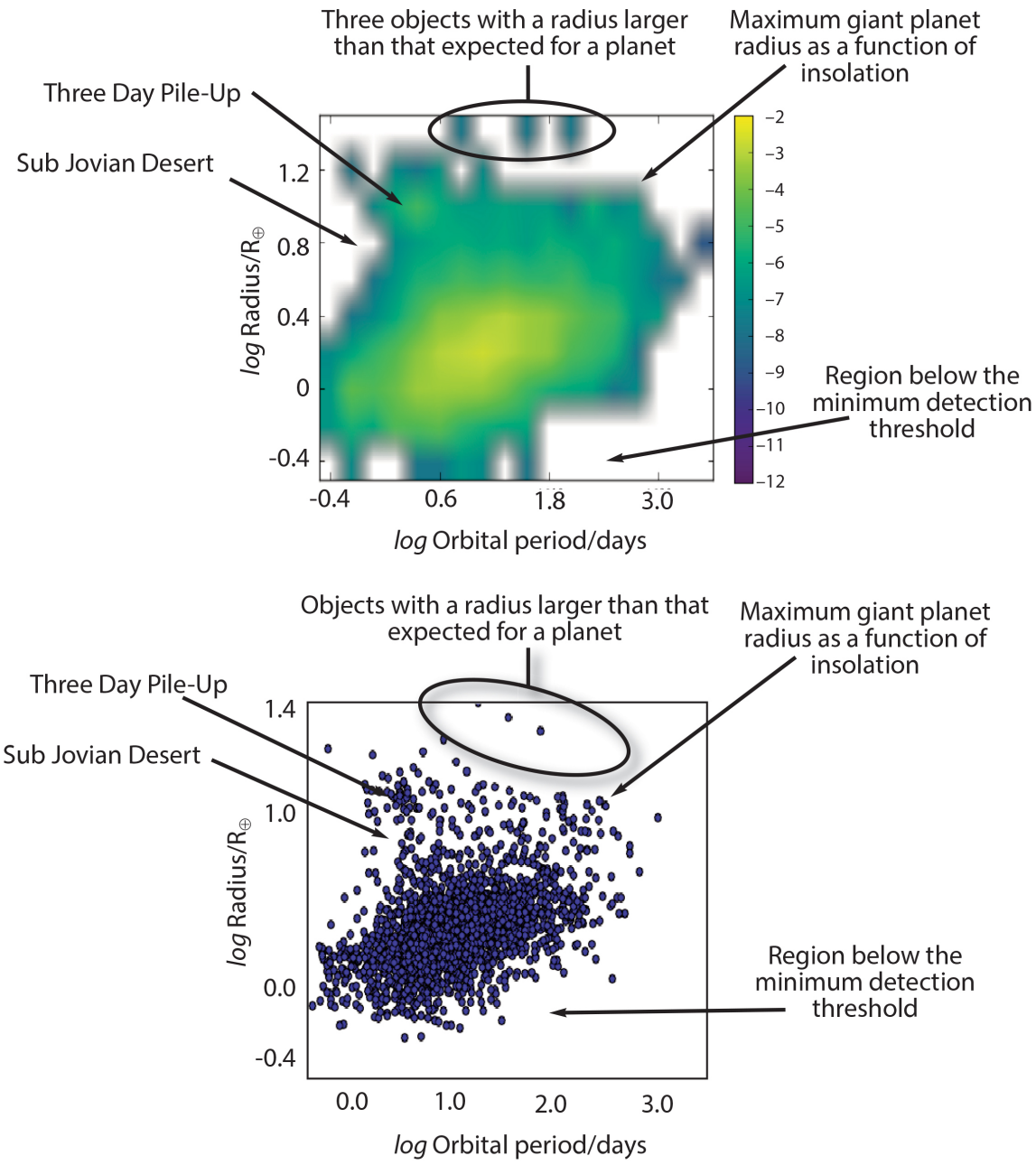


Fig. 6.11 Annotated contour map and scatter plot of the observed exoplanet distribution, *Kepler* DR25. Scale bar is logarithmic, and data in the contour plot has been normalised. The axes and labels do not match precisely because of the different types of data: in the upper panel each square representing a radius/period bin is of equal size, whereas in the lower panel the points in the scatter plot are distributed logarithmically.

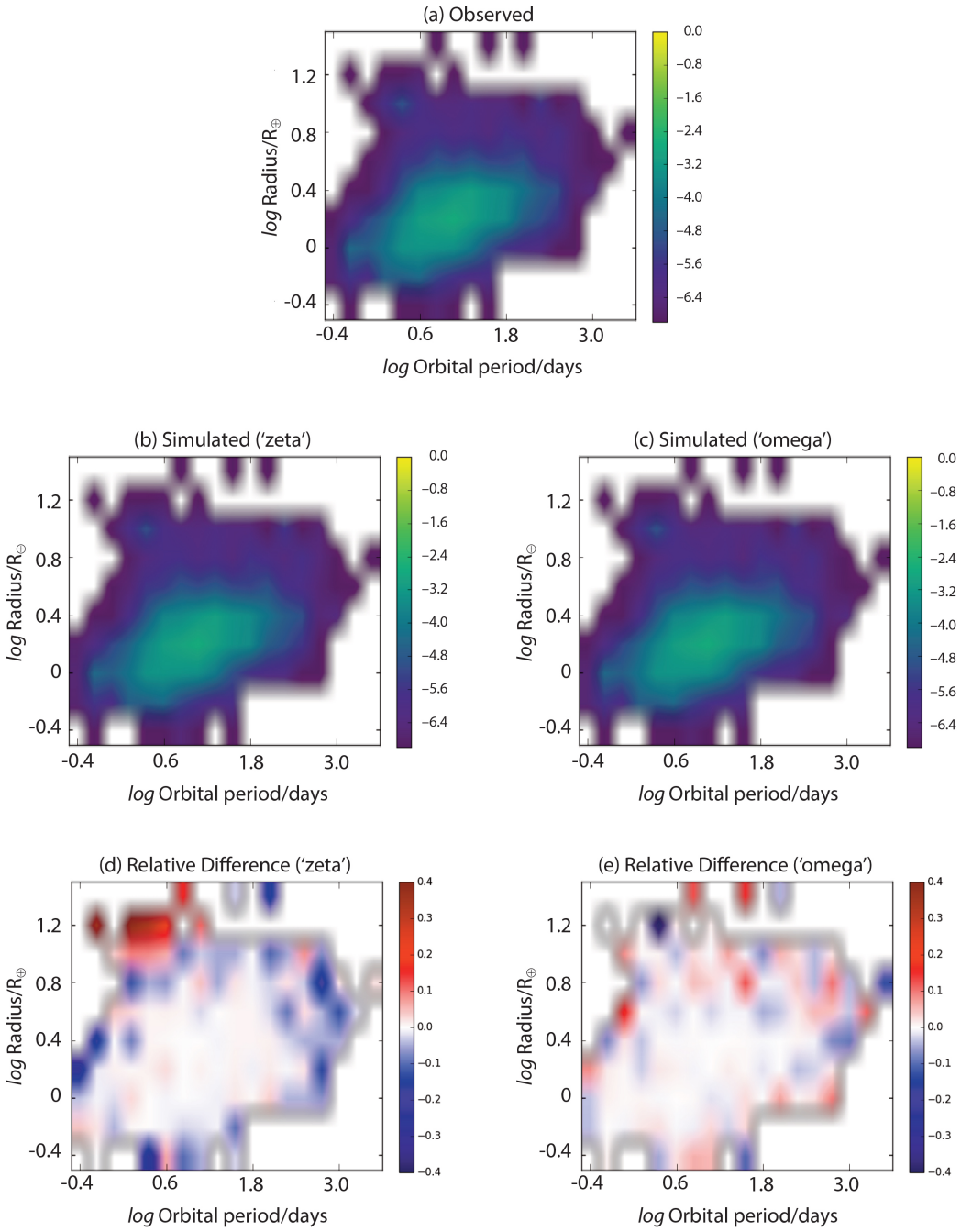


Fig. 6.12 Normalised log scale plots showing (a) confirmed planets, *Kepler* DR25, and simulated populations generated from (b) 'zeta' and (c) 'omega', in those bins which are populated in the observed population. (d) shows *log* confirmed planets minus *log* simulated population based on 'zeta' and (e) shows *log* confirmed planets minus *log* simulated population based on 'omega', populated bins only in both (d) and (e), linear scale in log space (so 0.10 indicates *log* 0.10, for example).

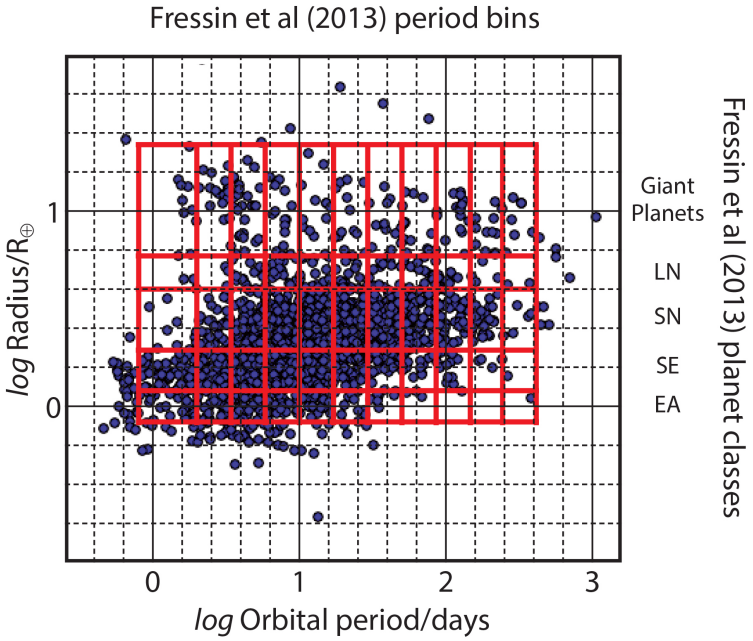


Fig. 6.13 Distribution of observed *Kepler* planets, *Kepler* DR25, radius against period. The background grid (black) indicates the logarithmic bins used in this work; the foreground grid (red) indicates the bins used in Fressin et al. (2013). LN = Large Neptunes, SN = Small Neptunes, SE = SuperEarths and EA = Earth-Analogues.

Table 6.12 Test exoplanet distribution, fields drawn from PLATO simulations, planets observable by *Kepler*. Main sequence single stars only, $T_{eff} < 7500$. The field observed by Kepler was $105^{\circ\,2}$, the field included in these test simulations was $194^{\circ\,2}$ as it included gaps between CCDs as well as some of the area surrounding the *Kepler* field.

	'zeta'	'omega'
Confirmed planets when data was accessed	2293	2293
Confirmed planets with estimated radii when data was accessed	2291	2291
Normalisation factor	15	1.5
Synthetic planets observable after blending, $Kp < 26$	4296	5004
As above, normalised to observed Kepler field, 105/194	2325	2708
Synthetic planets observable before blending, $Kp < 16$	4906	5513
As above, normalised to observed Kepler field, 105/194	2777	2984
Synthetic planets observable after blending, $Kp < 16$	4092	4350
As above, normalised to observed Kepler field, 105/194	2215	2354

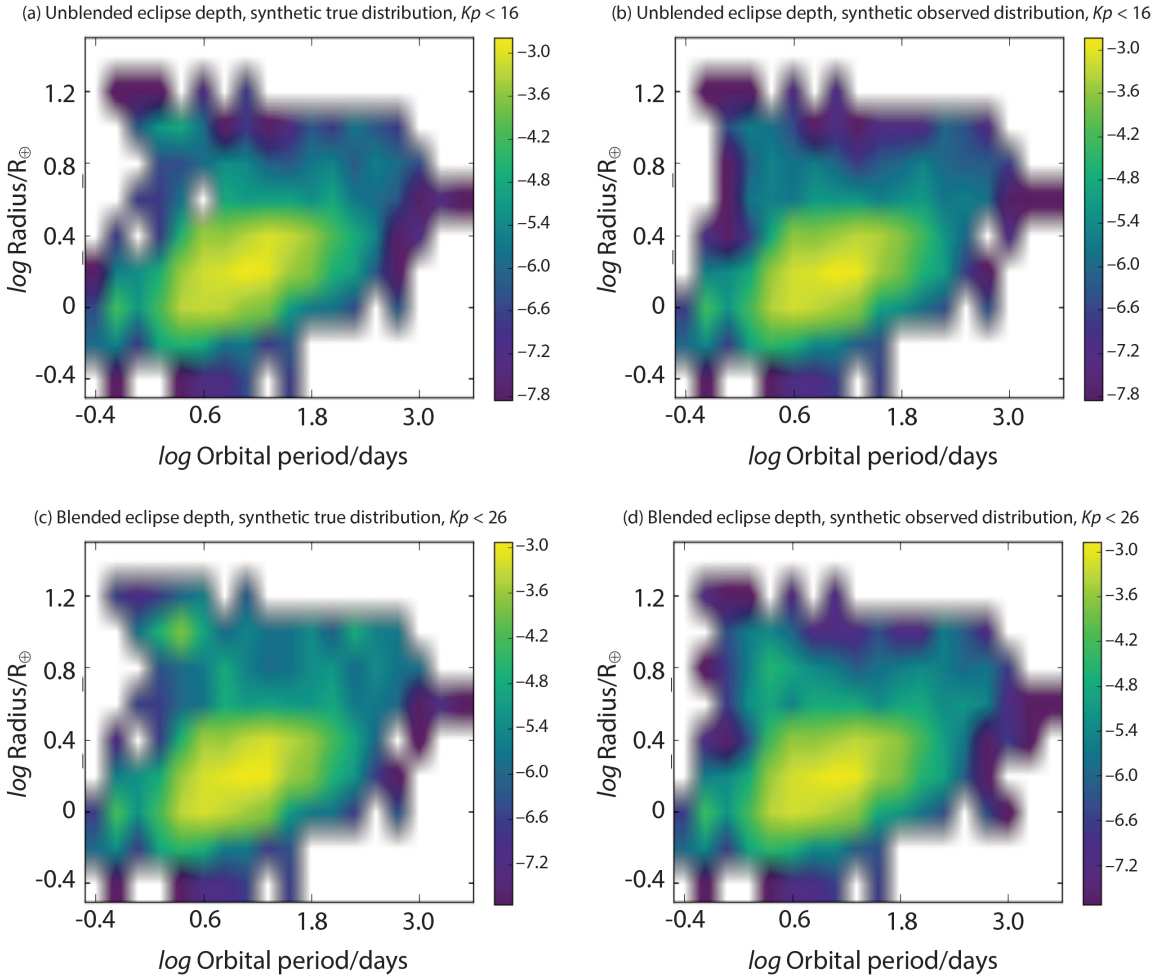


Fig. 6.14 Contour plots showing the results of the test output from the PLATO simulations, derived using the 'zeta' PDF, binned radius against binned period. (a) true planet radius, and (b) the apparent radius of the population in (a), once blending within the pixel is taken into account, $< Kp = 16$, $T_{eff} < 7500$ K. (c) the true radius of the planets observable after blending within the pixel and (d) the apparent radius of the population in (c), $Kp < 26$, $T_{eff} < 7500$ K. (a) and (c) show the actual planet radius, (b) and (d) show the observed planet radius, assuming the fractional transit depth is being evaluated against the true planet host. *Kepler* sensitivity in these plots is based on a 30 minute integration.

planets in the *Kepler* catalogue. This is as expected, as *Kepler* did not observe every star in the field. In the sample simulated with 'omega', more synthetic planets are considered detectable at the matching sensitivity than was the case with the sample simulated with 'zeta'. These simulations are part of the data set interpreted in Section 6.4 in the context of PLATO, where simulations based on 'omega' are shown to return significantly fewer terrestrial planets than simulations based on 'zeta', although numbers of giant planets are similar.

These results are based on simulations where the synthetic planet hosts were limited to main sequence single stars. Most *Kepler* planets orbit main sequence stars, although a few have been identified around red giants: for example, *Kepler*-91 b (Lillo-Box et al., 2014), which it is believed will be engulfed by its host within 55 M yr, and *Kepler*-432 b (Ortiz et al., 2015), which orbits an evolved star in an S-type orbit in a binary system. Such cases are, however, rare. Issues concerning the assumption of single star hosts were discussed in Section 6.2.1: however, 'zeta' and 'omega' were generated from data which was based on the single star assumption, so that assumption has been carried forward in this work.

The simulated number of planets reported in unpopulated bins of the observed planet sample is very small, and in many cases their presence can be explained through the reporting issues due to machine precision noted in Section 6.1.4. No planets are included in the PLATO simulations from the bins representing a planet radius of $1.4 < log\ R < 1.6$ $R_{\oplus}$, although these bins were included in deriving 'zeta' and 'omega', for reasons which are explored in the following sub-subsection, when the general characteristics of the distribution are discussed.

In Fig. 6.14, the assumptions are that the radius of the planet is measured against the radius of the planet host, that the radius of the planet host is correctly known and that the planet host is on the (synthetic) target list.

The differences in panels (a) and (b), Fig. 6.14, show how blending can remove planets from the observed distribution Panels (a) and (c) show the "sub Jovian desert", but in panels (b) and (d) this parameter space is occupied by blended hot Jupiters, whose true radii are larger than indicated in these panels. For reference, Fig. 6.11 indicates the location of the three-day pile-up and the sub-Jovian desert. Some long period/small radius planets will only be observable unblended, because when blended the precision of the observatory makes them unobservable. Table 6.12 indicates that in the test field simulated using 'zeta', at $Kp < 16$ 814 (17%) of the synthetic planets detectable unblended were not detectable once blending is taken into account, and in the test field simulated using 'omega', at $Kp < 16$ 1163 (21%) of the synthetic planets detectable unblended were not detectable once blending is taken into account.

To examine the question of dilution by blending as a function of magnitude, Fig 6.15 ('zeta' PDF) plots apparent magnitude against planet radius for the region covered in Fig. 6.14 and Table 6.12. The panels in Fig 6.15 correspond to the equivalent panels in Fig. 6.14. The

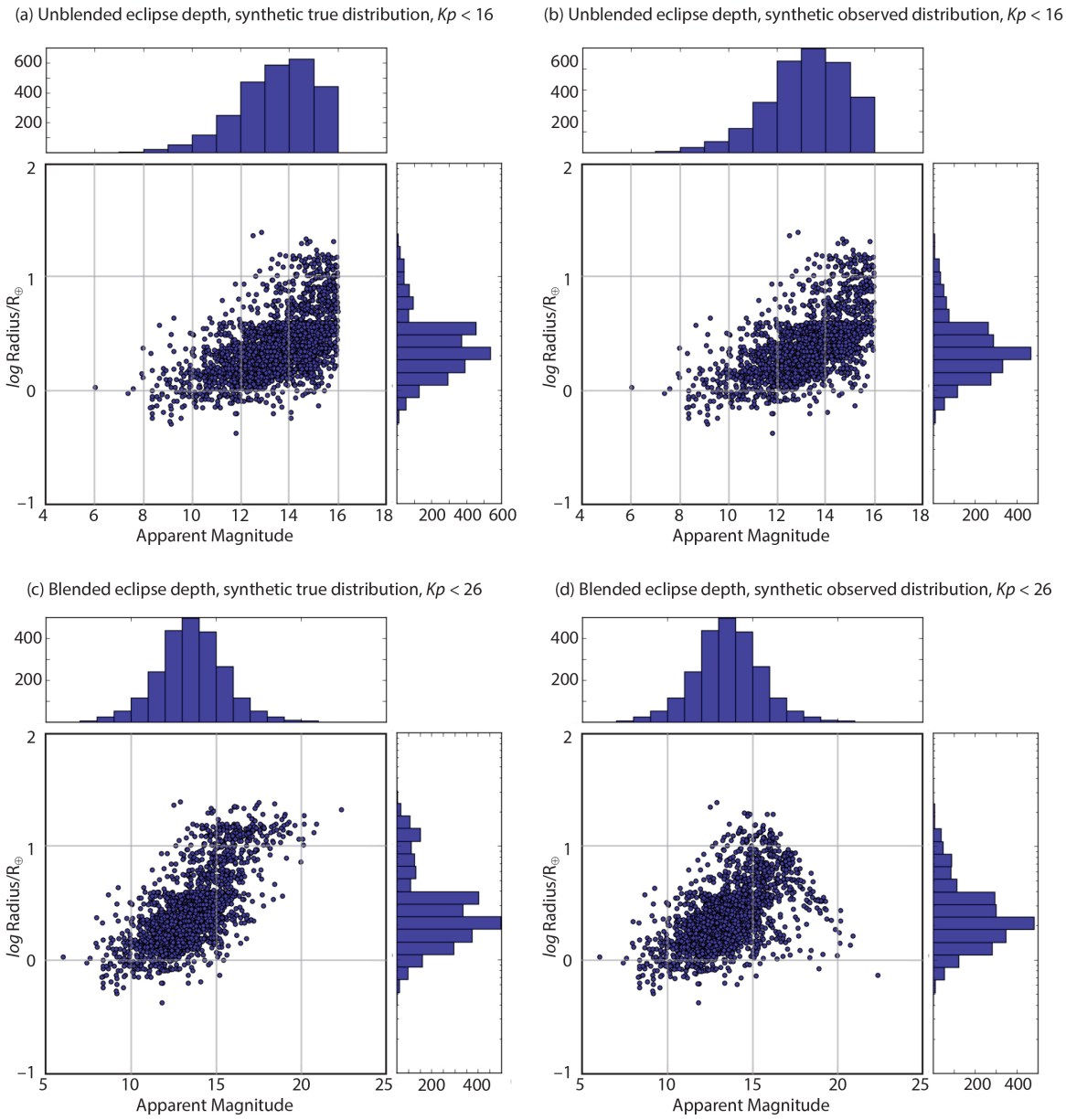


Fig. 6.15 Planet radius against apparent magnitude for the region covered in Fig. 6.14, 'zeta' PDF. (a) and (b): unblended, 30 minute integration, $Kp < 16$, $T_{eff} < 7500$ K. (c) and (d): allowing for blending, 30 minute integration, $Kp < 26$, $T_{eff} < 7500$ K. (a) and (c) show the actual planet radius, (b) and (d) show the observed planet radius once blending is taken into account.

faintest observable stars in (c) and (d) are $Kp \approx 23$, well within the $Kp < 26$ limit set by the simulation.

Comparing panels (c) and (d) in Fig 6.15, an interesting trend emerges at $Kp > 15$. In panel (c) (true radius) the faintest observable stars are giant planets, but in panel (d) (apparent radius) these same planets appear to be terrestrial, a direct consequence of dilution by blending. Stars fainter than $Kp = 16$ are unlikely to be targets in transiting exoplanet surveys, but may contaminate stars that are included.

6.3.1 General characteristics of the intrinsic exoplanet distribution

As discussed in Section 6.1, of the 180 period/radius bins, 69 contained no confirmed planets in *Kepler* DR25. Different physical conditions determine the reasons for the lack of planet detections in different regions. Fig. 6.11 illustrates three of these regions: the region below the minimum detection threshold, the sub-Jovian desert and the maximum giant planet radius as a function of insolation.

The minimum detection threshold is a consequence of the geometric transit probability, combined with *Kepler* sensitivity in a four year survey. For each observatory, there is a limit below which it is simply not possible to resolve a transit. The PDF below the minimum detection threshold describes an approximate upper limit on the intrinsic planet occurence rate, both in 'zeta' and 'omega'.

In other regions, noise is a consideration. This is particularly evident when considering the smallest planets (< 0.63 $R_\oplus$) at short periods (< 40 days).

For a given star, Neptune-like and giant planets are easier to detect at any orbital period than terrestrial planets as the transit depth will be deeper. Therefore, there should be planets detectable in the region identified in Fig. 6.11 as the sub-Jovian desert: that they are not observed in exoplanet surveys, including *Kepler*, indicates their absence has an astrophysical cause. The "three-day pile-up" (Wu and Lithwick (2011), Beaugé and Nesvorný (2012)), an excess of giant planets with $P \approx 3$ days, also identified in surveys including WASP and SuperWASP, is also apparent in *Kepler* data and is identified in Fig. 6.11.

Planets with estimated radii $R/R_\oplus > 40$ (log 1.6 $R_\oplus$) were excluded from the derivation of the PDFs. Planets in the range $25 < R/R_\oplus < 40$ ($1.4 < log\ R/R_\oplus < 1.6$) were included with caution in the calculations to derive the PDFs, but excluded from later simulations. Allowing for uncertainties, ≈ 25 $R_\oplus$ is a generally accepted maximum planet radius. Exoplanet.eu[6], which lists confirmed exoplanets from a wide variety of sources, lists 2780 planets as having been detected by primary transit (9 February 2018), the largest of which is HAT-P-67 b at 23.37 $R_\oplus$. Two planets detected by direct imaging are larger, with estimated radii of 28 $R_\oplus$

[6]http://http://exoplanet.eu/catalog/

(ROXs 42B b) and 25 $R_\oplus$ (CT Cha b) respectively, but all other planets from all detection methods for which there is an estimated radius are smaller than HAT-P-67 b. There are three bodies with estimated radii in the range $25 < R/R_\oplus < 40$ in DR25. There is the possibility that these objects are something other than planets, such as brown dwarfs. Fressin et al. (2013) excluded planets > 22 $R_\oplus$ from their analysis.

For planets in DR25 where the estimated radius is $10 < R/R_\oplus < 25$ and $P > 10$ days, maximum radius generally decreases as period increases (see Fig. 6.13). This is consistent with the concept that in a degenerate body such as a giant planet, maximum radius decreases as distance from the star increases due to reduced insolation (Mordasini et al., 2012). This feature is also annotated in Fig. 6.11.

6.3.2 Planet Classes

Throughout this subsection, the radius bins referred to can be identified with reference to Fig. 6.13, which shows the confirmed distribution overlayed by grid lines indicating the period and radius bins. The widely used classes from Fressin et al. (2013) are also indicated in Fig. 6.13. Additionally, new discoveries from ongoing analysis of *Kepler* data result in the number of confirmed planets on NExSci being updated from time to time. The figures below are the ones from the time the data was accessed and downloaded (9 May 2017). The breakdown of these figures is given in Table 6.13. Data from *Kepler* DR24 is included in Table 6.13 as well as data from *Kepler* DR25 to illustrate the effect that the updated stellar parameters had on the bulk planet radius distribution.

Mars-analogues: Mars has a radius of 0.532 $R_\oplus$ and hence would fall into the smallest radius bin, $-0.4 < log\ R/R_\oplus < -0.2$. Planets in this category have $log\ P < 1.6$ days (40 days) in DR25. One confirmed planet is smaller than $log\ -0.4$ $R_\oplus$ (0.40 $R_\oplus$) in *Kepler* DR25: this has been excluded from analysis.

Jackson et al. (2009) proposed that gaseous planets are tidally disrupted within 0.05 AU of the host star, and authors including Haswell et al. (2012), Bochinski et al. (2013) and Staab et al. (2017) have studied the clouds of circumstellar absorbing gas indicative of mass loss from larger planets, such as WASP-12 b. The Mars-analogue planets at very short orbital periods (< 1 day) may be the remains of tidally disrupted planets, or may represent giant or Neptune-like planets which have lost their H/He atmosphere through this or other means, leaving a dense, compressed core (Hébrard et al. (2004), Mocquet et al. (2014)).

Earth-analogues: The two radius bins $-0.2 < log\ R/R_\oplus < 0.0$ and $0.0 < log\ R/R_\oplus < 0.2$ cover the entire range designated by Fressin et al. (2013) as "Earth-analogues" ($0.8 < R/R_\oplus < 1.25$) with an overlap to planets < 0.8 $R_\oplus$ too small to be included in the Fressin et al. (2013)

Table 6.13 Confirmed planets, *Kepler* DR24 (12 May 2016) and *Kepler* DR25 (9 May 2017), by radius bin. Data from DR24 is included as a measure of the effect of the updated stellar radii used in DR25 in reducing the number of identified terrestrial planets.

Name	$R/R_\oplus$ *log*	$R/R_\oplus$ linear	DR24	DR25
Mars-analogue	-0.4 to -0.2	0.40 to 0.63	14 (0.61%)	10 (0.44%)
Earth-analogue 1	-0.2 to 0.0	0.63 to 1.0	153 (6.7%)	126 (5.5%)
Earth-analogue 2	0.0 to 0.2	1.0 to 1.6	558 (24.4%)	490 (21.4%)
Small Neptune 1	0.2 to 0.4	1.6 to 2.5	792 (34.6%)	797 (34.8%)
Small Neptune 2	0.4 to 0.6	2.5 to 4.0	506 (22.1%)	584 (25.5%)
Large Neptune	0.6 to 0.8	4.0 to 6.3	121 (5.3%)	133 (5.8%)
Giant Planets 1	0.8 to 1.0	6.3 to 10	63 (2.8%)	67 (2.8%)
Giant Planets 2	1.0 to 1.2	10 to 16	57 (2.5%)	68 (3.0%)
Giant Planets 3	1.2 to 1.4	16 to 25	10 (0.44%)	9 (0.39%)
Brown dwarfs?	1.4 to 1.6	25 to 40	8 (0.35%)	3 (0.13%)

sample, and planets > 1.25 $R_\oplus$ which are included in the Fressin et al. (2013) SuperEarth category.

As with *Kepler*, a key aim for PLATO is detecting a twin to Earth. For *Kepler*, as shown in Fig. 6.13, a 1 $R_\oplus$ planet in a one year orbit around a 1 $R_\odot$ star was right on the limit of detectability, and so it is not possible from the *Kepler* data alone to make a firm prediction of the number of Earth twins PLATO will observe.

Small Neptunes: The two radius bins $0.2 < log\ R/R_\oplus < 0.4$ and $0.4 < log\ R/R_\oplus < 0.6$ cover the entire range designated by Fressin et al. (2013) as "Small Neptunes" ($2 < R/R_\oplus < 4$) along with the remainder of the "superEarths" not included in our "Earth-analogue" class. According to the NExSci cumulative table this is the most populated region. The orbital period is $-0.2 < log\ P/\text{days} < 2.8$. The absence of planets of this radius at shorter periods forms part of the sub-Jovian desert.

Stellar irradiation on a H/He atmosphere will cause that atmosphere to expand, and will affect gaseous planets such as those in this class much more than smaller rocky planets. However, studies of the sub-Jovian desert refer to an absence of planets at short periods between Earth- and Jupiter-***mass*** as well as radius: so insolation alone cannot provide the answer. As Fig. 6.13 and Fig. 6.11 show, the minimum orbital period increases with planet radius, with rare exceptions, up to $\approx$ *log* 0.8 $R_\oplus$ and *log* 0.2 days, where the well studied "three day pile-up" is found. This is despite the fact that larger exoplanets should be easier to detect at a given distance from the host star than smaller ones.

Large Neptunes: The radius bin $0.6 < log\ R/R_\oplus < 0.8$ covers the entire range designated by Fressin et al. (2013) as "Large Neptunes" ($4 < R/R_\oplus < 6$) along with some "Giant Planets". Whether binning planet radius logarithmically or using the classes from Fressin et al. (2013), this bin represents a significant drop in numbers on the previous bin. This is consistent with theories on planet formation which indicate that smaller planets are more common than larger ones It also justifies the division by Fressin et al. (2013) of a single Neptune bin used by earlier authors into a "Small Neptune" and "Large Neptune" bin in their work.

Giant Planets: The radius bins $0.8 < log\ R/R_\oplus < 1.0$, $1.0 < log\ R/R_\oplus < 1.2$ and $1.2 < log\ R/R_\oplus < 1.4$ cover most of the range designated by Fressin et al. (2013) as "giant planets" ($6 < R/R_\oplus$ 22) along with some planets too large to be classified by Fressin et al. (2013). The "three day pile-up", an excess of hot Jupiters at about 3 days, is evident in the observed population, Fig. 6.11, and is also reflected in the synthetic population.

Giant planets were the one class in which Fressin et al. (2013) noticed a mismatch between observations and simulations. As they report, small stars are believed not to host large planets. In this work, a similar mismatch is apparent when unblended depths are used with *Kepler* precision. However, when allowing for blending, the mismatch appears to disappear, with giant planets appearing to be Neptune-like or terrestrial following blending. Given that giant planets will not be selected by PLATO for follow up and because, as shown in the following section, M-class stars observable by PLATO hosting giant planets are few and far between, this mismatch has not been corrected for at this stage, although this is a consideration for future work. Correcting for blending in unresolvable binaries may well account for the mismatch: further study will confirm or refute this idea.

Brown dwarfs?: The radius bin $1.4 < log\ R/R_\oplus < 1.6$ covers objects designated as planets, but with an estimated radius larger than would be expected of a planet. Either the radius has been overestimated, or these are other astrophysical objects such as brown dwarfs. There are also three larger confirmed planets on NExSci: if the radius estimate is correct, these are also too large to be planets.

6.4 Application to the PLATO fields

The purpose of obtaining the intrinsic exoplanet distribution was to apply it to the PLATO fields, to obtain an estimate of the ratio of unblended planets to both blended & unblended eclipsing binaries, and to quantify the effect of blending within the pixel in diluting the apparent planet radius.

To that end, the intrinsic exoplanet distributions were applied to two full synthetic PLATO fields, one centred at $l = 65°$, $b = 30°$, the proposed Long Look North (LLN) field, and the other centred on $l = 253°$, $b = -30°$, the proposed Long Look South (LLS) field. The

synthetic LLN field includes within it the region discussed in previous sections for testing and verification purposes.

The synthetic populations were analysed taking into account the anticipated performance of PLATO in a one hour integration for six, seven or eight cameras/group (the original proposal was for eight cameras/group and the mission received final approval with six cameras/group), and for the PLATO priority populations, described in Table 1.6. The code provided by Claude Catala to determine the PLATO precision and overlap was utilised in combination with updated signal and noise estimates from the Instrument Noise Budget[7] (Chapter 5). The *V* magnitude band has been used, as this is the band used to define the PLATO priority populations.

As indicated by Eq. 5.1 and Eq. 5.12, the number of events observed is key in determining the transit SNR or, as it was known with *Kepler*, the multiple event statistic. In the same way that Table 6.4 considered transit SNR as a function of period and magnitude for *Kepler*, so Table 6.14 considers the maximum orbital period, in the context of our binning, that a planet orbiting a Sun-like star could achieve a transit SNR ≥ 7.1 in a one, two or three year PLATO observation, where the rms photometric precision is equal to the target for PLATO for the central field for the P5 population, namely 80 ppm. Again, I consider four planets: Jupiter, Neptune, a 2 $R_{\oplus}$ planet and Earth. The period at which Earth would achieve the required transit SNR may appear short, but remember that at a transit depth of 84 ppm the transit SNR for a single event for Earth where rms photometric precision is 80 ppm is close to unity. The calculated intrinsic exoplanet distribution in the *log* R-*log* P bins where a MES of 7.1 for *Kepler* would not be achieved in a four year observation is already low, due to the requirement to match unpopulated bins, so few planets were simulated in these bins in the PLATO simulations. Planets where the transit SNR for a single event is less than unity when considered against the magnitude and surface temperature of the host star in question would also not be considered detectable in my simulations, even though Eq. 5.1 and Eq. 5.12 imply that phase folding may make such planets detectable if a sufficient number of events were analysed. Fig. 6.1 indicates that a proportion of stars are intrinsically quieter than a rms photometric precision calculation may indicate, allowing shallower transits to be detected at a given magnitude than might theoretically be expected. PLATO is prioritising F5-K stars, and Martínez-Arnáiz et al. (2010) find that G class stars are quieter on average than other stars. Therefore, PLATO can be expected to observe planets around quiet stars, making the planets more likely to be detected than would otherwise be expected as a function of magnitude and period. Few planets that would have a transit SNR < 7.1 as a function of period, magnitude and surface temperature are identified in my simulations, and those that are have been allowed to remain as representatives of planets around quiet stars.

[7]PLATO-DLR-PL-RP-001, issue 3, 19 Feb 2016

Table 6.14 Theoretical detectability of a PLATO planet with a transit SNR $\geq$ 7.1 at rms photometric noise at the target level for P5 in the central part of the field, 80 ppm. Data as presented relates to the period bins used in this work, and is max log P_{orb} where transit SNR $/ge$ 7.1. Four planets are considered against a Sun-like host: Jupiter (transit depth 10100 ppm), Neptune (transit depth 1250 ppm), a 2 $R_\oplus$ planet (transit depth 336 ppm) and Earth (transit depth 84 ppm). One, two and three year observations are considered. Note that the transit SNR for an Earth-sized planet in a single observation at this photometric precision is close to unity.

Period of observations	Jupiter log *P*/days	Neptune log *P*/days	2 $R_\oplus$ planet log *P*/days	Earth
1 year	$>$ 3.2 days	$>$ 3.2 days	2.2 days	1.0 days
2 years	$>$ 3.2 days	$>$ 3.2 days	2.6 days	1.2 days
3 years	$>$ 3.2 days	$>$ 3.2 days	2.8 days	1.4 days

Ten full field simulations were obtained in each global field with the intrinsic exoplanet distributions 'zeta' and 'omega'. The mean of these ten simulations is reported in this section.

6.4.1 Overview

The total numbers of planets are given in Table 6.15 for simulations based on the 'zeta' and 'omega' populations. These are given by priority population, by number of cameras and by whether or not possible blending with stars in the same pixel is taken into account. Also included are all systems with stellar class F5-M, $V < 16$, a population analogous to that observed by *Kepler*, and all planets detectable by PLATO, that is all stellar classes, $V< 26$. P2 is a subset of P1, in the same way that P1 is a subset of P5 (Table 1.6). The difference between the F5-M, $V < 16$ systems and the sum of P4 and P5 is the stars in stellar class F5-K7 of $13 < V < 16$.

Stars selected for the PLATO exoplanet simulations presented here are all MS single stars.

My estimates of the number of terrestrial planets PLATO is likely to identify, depend critically on the true completeness and sensitivity of *Kepler*. The benchmark estimate in the *Kepler* Instrument Handbook is more sensitive than the standard set for PLATO in a one hour integration, which in turn is more sensitive than a single *Kepler* long cadence. In the PLATO priority classes, 'zeta' returns an order of magnitude more terrestrial planets observable by PLATO than does 'omega', a natural consequence of these different sensitivities.

More synthetic planets are considered detectable by PLATO in the LLS field than the LLN field, despite the fact that LLS covers an area further from the Galactic centre. This is believed to be due to the fact that to observe the LLS field, it is necessary to look through the Galactic plane. This finding applies to synthetic populations based on both 'zeta' and

Table 6.15 Synthetic exoplanets, PLATO LLN and LLS, total numbers. Error is σ. Mean of 10 simulations, one hour integration, using the PDFs 'zeta' and 'omega'.

Population	32 cameras Unblended	32 cameras Inc Blending	28 cameras Unblended	28 cameras Inc Blending	24 cameras Unblended	24 cameras Inc Blending
$V < 26$						
LLN 'zeta'	$3.01 \pm 0.02 \times 10^4$	$2.50 \pm 0.02 \times 10^4$	$2.75 \pm 0.02 \times 10^4$	$2.28 \pm 0.02 \times 10^4$	$2.47 \pm 0.02 \times 10^4$	$2.05 \pm 0.02 \times 10^4$
LLN 'omega'	$1.12 \pm 0.01 \times 10^4$	$8.32 \pm 0.06 \times 10^3$	$1.05 \pm 0.01 \times 10^4$	$7.79 \pm 0.05 \times 10^3$	$9.65 \pm 0.07 \times 10^3$	$7.20 \pm 0.05 \times 10^3$
LLS 'zeta'	$3.07 \pm 0.02 \times 10^4$	$2.71 \pm 0.02 \times 10^4$	$2.80 \pm 0.02 \times 10^4$	$2.48 \pm 0.02 \times 10^4$	$2.52 \pm 0.01 \times 10^4$	$2.23 \pm 0.02 \times 10^4$
LLS 'omega'	$1.04 \pm 0.06 \times 10^4$	$8.61 \pm 0.05 \times 10^3$	$9.73 \pm 0.06 \times 10^3$	$8.06 \pm 0.05 \times 10^3$	$8.99 \pm 0.06 \times 10^3$	$7.46 \pm 0.06 \times 10^3$
$V < 16$, F5-M						
LLN 'zeta'	$2.26 \pm 0.02 \times 10^4$	$2.04 \pm 0.02 \times 10^4$	$2.07 \pm 0.02 \times 10^4$	$1.86 \pm 0.02 \times 10^4$	$1.86 \pm 0.02 \times 10^4$	$1.68 \pm 0.02 \times 10^4$
LLN 'omega'	$7.11 \pm 0.08 \times 10^3$	$6.29 \pm 0.07 \times 10^3$	$6.66 \pm 0.08 \times 10^3$	$5.90 \pm 0.05 \times 10^3$	$6.18 \pm 0.07 \times 10^3$	$5.46 \pm 0.06 \times 10^3$
LLS 'zeta'	$2.44 \pm 0.02 \times 10^4$	$2.25 \pm 0.02 \times 10^4$	$2.23 \pm 0.01 \times 10^4$	$2.06 \pm 0.01 \times 10^4$	$2.00 \pm 0.01 \times 10^4$	$1.86 \pm 0.01 \times 10^4$
LLS 'omega'	$7.33 \pm 0.02 \times 10^3$	$6.71 \pm 0.02 \times 10^3$	$6.87 \pm 0.02 \times 10^3$	$6.29 \pm 0.03 \times 10^3$	$6.36 \pm 0.02 \times 10^3$	$5.84 \pm 0.03 \times 10^3$
P1						
LLN 'zeta'	$1.95 \pm 0.05 \times 10^3$	$1.93 \pm 0.05 \times 10^3$	$1.83 \pm 0.05 \times 10^3$	$1.82 \pm 0.05 \times 10^3$	$1.70 \pm 0.05 \times 10^3$	$1.69 \pm 0.05 \times 10^3$
LLN 'omega'	270 ± 10	270 ± 10	261 ± 9	260 ± 9	250 ± 10	250 ± 10
LLS 'zeta'	$2.27 \pm 0.06 \times 10^3$	$2.25 \pm 0.06 \times 10^3$	$2.12 \pm 0.06 \times 10^3$	$2.11 \pm 0.06 \times 10^3$	$1.98 \pm 0.06 \times 10^3$	$1.97 \pm 0.06 \times 10^3$
LLS 'omega'	331 ± 9	330 ± 9	319 ± 9	318 ± 10	305 ± 8	304 ± 8
P2						
LLN 'zeta'	66 ± 10	66 ± 10	63 ± 10	63 ± 10	61 ± 10	61 ± 10
LLN 'omega'	7 ± 2	7 ± 2	7 ± 2	7 ± 2	6 ± 2	6 ± 2
LLS 'zeta'	76 ± 12	76 ± 12	73 ± 12	73 ± 12	71 ± 11	71 ± 11
LLS 'omega'	6 ± 1	6 ± 1	6 ± 1	6 ± 1	6 ± 1	6 ± 1
P4						
LLN 'zeta'	$1.11 \pm 0.02 \times 10^3$	$0.98 \pm 0.02 \times 10^3$	$1.04 \pm 0.02 \times 10^3$	$0.91 \pm 0.02 \times 10^3$	$0.97 \pm 0.02 \times 10^3$	$0.85 \pm 0.04 \times 10^3$
LLN 'omega'	180 ± 10	160 ± 10	170 ± 10	150 ± 10	170 ± 10	150 ± 10
LLS 'zeta'	$1.27 \pm 0.03 \times 10^3$	$1.15 \pm 0.03 \times 10^3$	$1.19 \pm 0.03 \times 10^3$	$1.07 \pm 0.03 \times 10^3$	$1.10 \pm 0.03 \times 10^3$	$0.99 \pm 0.03 \times 10^3$
LLS 'omega'	200 ± 10	185 ± 9	200 ± 10	179 ± 8	180 ± 10	170 ± 9
P5						
LLN 'zeta'	$8.6 \pm 0.1 \times 10^3$	$8.2 \pm 0.2 \times 10^3$	$7.8 \pm 0.1 \times 10^3$	$7.6 \pm 0.1 \times 10^3$	$7.1 \pm 0.1 \times 10^3$	$6.9 \pm 0.1 \times 10^3$
LLN 'omega'	$1.67 \pm 0.01 \times 10^3$	$1.63 \pm 0.01 \times 10^3$	$1.59 \pm 0.02 \times 10^3$	$1.55 \pm 0.01 \times 10^3$	$1.50 \pm 0.02 \times 10^3$	$1.46 \pm 0.02 \times 10^3$
LLS 'zeta'	$9.6 \pm 0.1 \times 10^3$	$9.4 \pm 0.1 \times 10^3$	$8.9 \pm 0.1 \times 10^3$	$8.67 \pm 0.09 \times 10^3$	$8.11 \pm 0.09 \times 10^3$	$7.92 \pm 0.09 \times 10^3$
LLS 'omega'	$1.88 \pm 0.02 \times 10^3$	$1.85 \pm 0.02 \times 10^3$	$1.79 \pm 0.02 \times 10^3$	$1.76 \pm 0.02 \times 10^3$	$1.69 \pm 0.02 \times 10^3$	1.66 ± 0.02

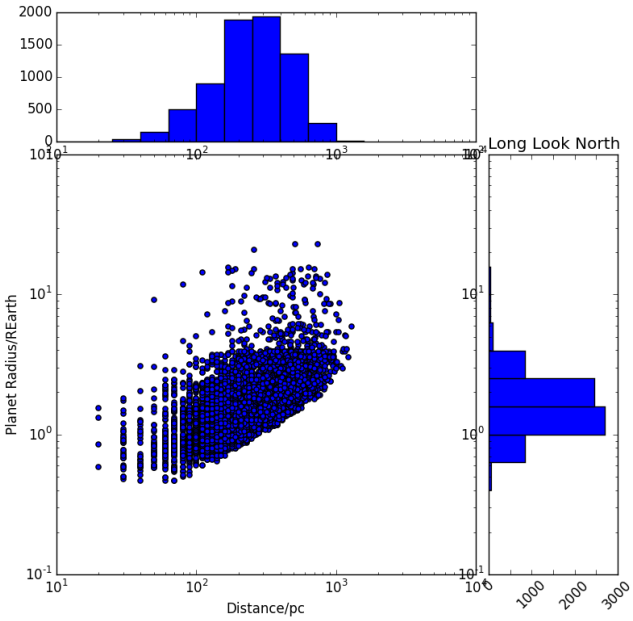


Fig. 6.16 A synthetic planet population extracted with PLATO precision, P5, plotting planet radius against distance from Earth. There is a clear relation between planet radius and the maximum distance at which it can be observed, which is a function of T_{eff} and apparent magnitude V. For comparison with *Kepler* limits (30 minute integration), see Fig. 6.5.

'omega'. That there is a real difference in the distribution of the proportion of synthetic single stars as a function of magnitude in the synthetic LLN and LLS fields is illustrated in Fig. 4.3 and Fig. 4.4, and discussed in Section 4.2. The proportion of both unblended and blended binaries in the two fields is discussed in Section 7.4. See also the discussion in Section 8.2.2 on the integrated stellar density along the line of sight to both LLN and LLS, as derived from the double exponential from which the thin and thick discs are simulated (Chapter 2), and Table E.13 to Table E.30, in Appendix E, which compare 1° x 1° regions in LLN and LLS by the numbers of stars simulated, once extinction is taken into account.

Comparison of the synthetic fields with Gaia DR2, as part of future work, will determine if this trend matches observations, or if it is an artifact of the simulations.

Just as Fig. 6.5 illustrates the relationship between planet radius and distance at which that planet can be observed by *Kepler* in a 30 minute integration, so Fig. 6.16 does the same for the PLATO P5 population with PLATO precision and a one hour integration (24 cameras, LLN), with Fig. 6.17 showing the same information for LLS. The two PLATO fields do not show any significant differences. The limit at which a planet of a given radius can be detected is a function of T_{eff} and apparent magnitude V, and is independent of which PDF, 'zeta' or 'omega', was used to simulate the population. Table E.28 to Table E.30 in Appendix E indicate that stars in the P5 priority population are not detectable beyond 2 kpc.

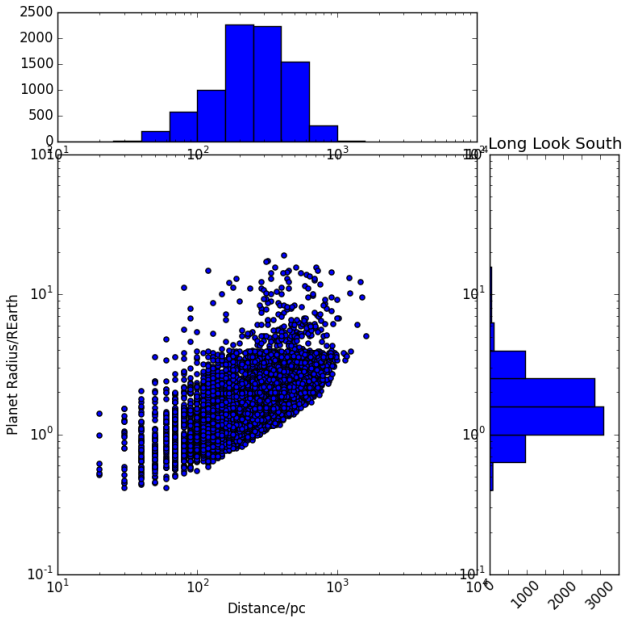


Fig. 6.17 As Fig. 6.16, Long Look South.

Fig. 6.18 plots the information from Fig. 6.16 and Fig. 6.17 in relation to the scale height of the Galaxy, as defined in Chapter 2. The position of the solar system, 30 pc above the Galactic plane, is indicated by the yellow dot. A magenta oval indicates the observational limit of 1 $R_\oplus$ planets, $\approx$ 300 pc at $b = |5.5|^\circ$ and $\approx$ 175 pc at $b = |50|^\circ$, while a cyan oval indicates the observational limit of 2 $R_\oplus$ planets, $\approx$ 1000 pc at $b = |5.5|^\circ$ and $\approx$ 400 pc at $b = |50|^\circ$. Double headed arrows indicate lines of sight $b = |15|^\circ$, $|30|^\circ$ and $|45|^\circ$. This indicates that Earth-analogue planets will be detected well within one vertical scale height of the Galactic plane in a one hour integration, and most larger planets within two scale heights in a one hour integration. The scale height is defined in Eq. 2.1.

The maximum detectable distance as a function of b was obtained by generating a series of plots similar to Fig. 6.16 and Fig. 6.17, with each plot limited to planets at $\pm$ 5° of b = 10°, 20°, 30°, 40° and 50° in LLN and $b = -10^\circ$, -20°, -30°, -40° and -50° in LLS. The maximum detectable distance of the 1 $R_\oplus$ and 2 $R_\oplus$ planets was used to derive the parameters of the ovals in Fig. 6.18. Again, a one hour integration was used.

Intuitively, as LLN is centred on l = 65° and LLS is centred on l = 253°, it would be expected that there would be more stellar systems in the LLN than LLS. The total number of synthetic systems with Gaia magnitudes in the range $4 < G < 26$ and $4 < G < 16$ in a 1° wide vertical strip in the centre of each field was assessed, as part of the process of calibrating the fields against data from Gaia DR1 and MAST described in Section 2.3. The results at $4 < G < 26$ for the strip in LLS were 4,100,000 and 77,000 stellar systems respectively, and

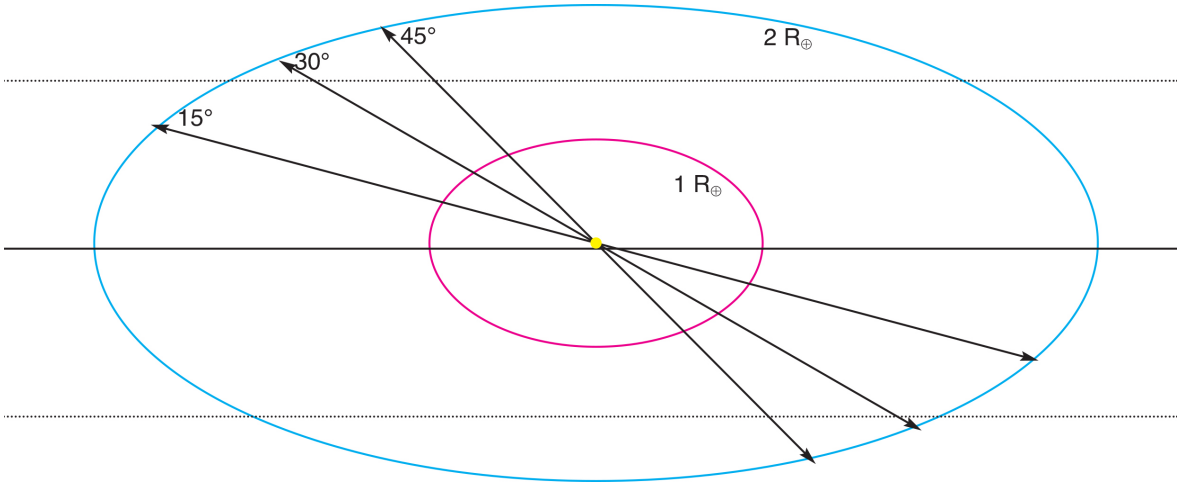


Fig. 6.18 Distance at which planets in the P5 population can be observed by PLATO. Solid line: Galactic plane. Dashed lines: one scale height of the thin disc (Eq. 2.1). Magenta oval: distance within which planets of 1 $R_\oplus$ can be observed by PLATO in a one hour integration. Cyan oval: distance within which planets of 2 $R_\oplus$ can be observed in a one hour integration.

for LLN were 8,700,000 and 110,000 stellar systems respectively, as expected. As shown in Fig. 4.3 and Fig. 4.4, the single star fraction for $4 < G < 26$ is higher in LLN, but for $4 < G < 16$ is higher in LLS. There are 5,100,000 single stars in the strip in LLN with $4 < G < 26$, and 51,000 with $4 < G < 16$, while in the strip in LLS there are 2,200,000 single stars with $4 < G < 26$ and 38,000 with $4 < G < 16$. So there are more planets detectable by PLATO in LLS, but more stellar systems, and more single systems, in LLN. The method used to simulate planets had seeded planets around all single stars in a given field, irrespective of distance, and discarded only those where the geometric transit probability made a transit detection impossible, whatever the sensitivity of the observatory, or where the planet would be inside the star. Therefore, more planets in total were simulated in LLN. It is only when the simulated population is interpreted with PLATO precision, as defined by the work described in Chapter 5, that there are more planets in LLS than LLN: so a greater proportion of planets in LLN are beyond the detection threshold. Therefore, in these simulations, there must be more nearby systems when looking through the Galactic plane than simply looking above it. That there is local variation is illustrated by the data in Table E.13 to Table E.30: however, over a large enough area, local differences appear to be smoothed out.

These results should be treated with caution, as my simulations treat the Galaxy as a homogeneous disc and the Galaxy is not homogeneous, and the extinction model may not have the required resolution. Testing has been completed in limited areas: more extensive testing against observations to confirm the best regions of the sky to observe are a subject of further work, especially once Gaia DR2 and DR3 are released. Both catalogues are expected

Table 6.16 Guide to PLATO synthetic exoplanet population tables and figures, PLATO simulations, LLN and LLS.

Population	Binning	True radius	Apparent radius	Table	Figure	Panel
$V < 26$	log	X		C.2	6.19	a
$V < 26$	log		X	C.2	6.20	a
$V < 26$	Fressin et al. (2013)	X		C.8	6.21	a
$V < 26$	Fressin et al. (2013)		X	C.8	6.22	a
$V < 16$, F5-M	log	X		C.3	6.19	b
$V < 16$, F5-M	log		X	C.3	6.20	b
$V < 16$, F5-M	Fressin et al. (2013)	X		C.9	6.21	b
$V < 16$, F5-M	Fressin et al. (2013)		X	C.9	6.22	b
P1	log	X		C.4	6.19	c
P1	log		X	C.4	6.20	c
P1	Fressin et al. (2013)	X		C.10	6.21	c
P1	Fressin et al. (2013)		X	C.10	6.22	c
P2	log	X		C.5	6.19	d
P2	log		X	C.5	6.20	d
P2	Fressin et al. (2013)	X		C.11	6.21	d
P2	Fressin et al. (2013)		X	C.11	6.22	d
P4	log	X		C.6	6.19	e
P4	log		X	C.6	6.20	e
P4	Fressin et al. (2013)	X		C.12	6.21	e
P4	Fressin et al. (2013)		X	C.12	6.22	e
P5	log		X	C.7	6.20	f
P5	log	X		C.7	6.19	f
P5	Fressin et al. (2013)	X		C.13	6.21	f
P5	Fressin et al. (2013)		X	C.13	6.22	f

to be more extensive than Gaia DR1. However, as a principle, in terms of overall numbers, it does seem that the LLS field is likely to produce a greater PLATO planet haul than LLN.

6.4.2 PLATO synthetic population by planet radius

The PLATO observing strategy (Rauer et al. (2014), PLATO Definition Study Report) is likely to adopt two long look fields, observed either for two years for each, or for three years for one (LLS) and one year for the other (LLN). The prediction shown in Rauer et al. (2014), comparing the expected haul of planets $< 2\ R_\oplus$ for PLATO to *Kepler* (Fressin et al., 2013) and TESS, predates Morton et al. (2016), which significantly expanded the *Kepler* planet haul. From my simulations, the key prediction from Rauer et al. (2014), that *the lifetime planet haul from PLATO of planets which may be fully characterised through asteroseismology of the host star or RV follow up will significantly exceed other surveys*, holds.

In the following I present, in Fig. 6.19 to Fig. 6.22 and in Table C.2 to Table C.13, the radius distribution of different synthetic planet samples, referring to different priority populations, as summarised in Table 6.16. The information is presented both in logarithmic bins in the *log R-log P* plane and in bins from Fressin et al. (2013). Where bins from Fressin

et al. (2013) are used, additional bins have been added to cover planets too large or too small to be considered by Fressin et al. (2013), but which are included within my simulations.

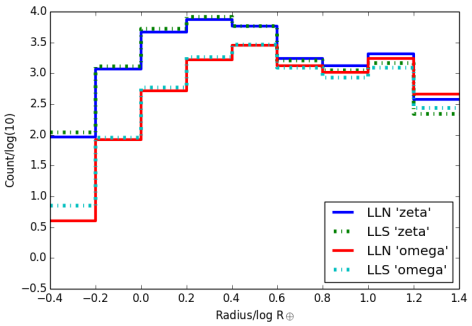
The true radii of the planets are considered, as are the apparent radii of the planets once blending within the pixel has been taken into account.

The relevant samples of confirmed planets from *Kepler* DR25 (April 2018) are included in Table C.2 to Table C.13 to compare the peaks in distribution by planet class as well as numbers of planets. These indicate that simulations based on the 'zeta' PDF, when interpreted by a PLATO one hour integration, often have a maximum in the same place as the equivalent observed *Kepler* population, whether the data is binned logarithmically or by bins from Fressin et al. (2013). The exception is P5 when using bins from Fressin et al. (2013), where the observed *Kepler* population peaks in the "Small Neptune" class ($2 < R/R_{\oplus} < 4$) while the synthetic population peaks in the "SuperEarth" class ($1.25 < R/R_{\oplus} < 2$). Where the synthetic population has been derived using 'omega', in bins from Fressin et al. (2013) the synthetic and observed populations generally have maxima in the same radius bins, with some exceptions in P4. Where data is binned logarithmically, the maxima in the population synthesised with 'omega' are consistently at a larger planet radius than the population synthesised with 'zeta', and in the P4, P5, $V < 26$ and $V < 16$, F5-M populations, are also systematically at larger radii than in the observed *Kepler* population. There are no confirmed *Kepler* planet hosts that would fall in P2, and very few that would fall in P1. *Kepler* pixels saturate at $Kp \leq 11$.

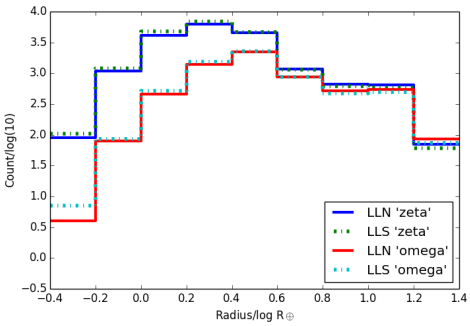
P2 is a subset of the P1 population and P1 is a subset of P5. The figures in Table 6.15 and the tables in Appendix C indicate that, for P1, P2 and P5 populations, the number of cameras in a group is of more significance in terms of numbers of simulated planets that can be detected than the effects of blending within a pixel, while for P4, which has a fainter magnitude limit, both the number of cameras and the inclusion of blending is significant.

If the data in *Kepler* DR25 is unaffected by dilution through blending through unrecognised stellar multiplicity and unresolved background contaminants, and if the same applies to PLATO observations, it appears that only the P4 PLATO priority population will be significantly affected by dilution by blending of the target star. If, as seems more likely, the *Kepler* DR25 data *is* affected by dilution through blending through unrecognised stellar multiplicity and unresolved background contaminants (Chapter 6.2.1), this will impact the distribution of confirmed planets on which the PDFs 'zeta' and 'omega' were based.

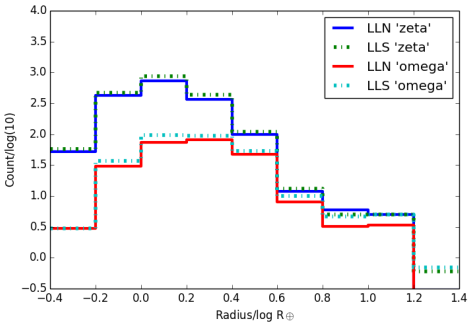
As shown in Fig. 6.19 to Fig. 6.20 and stated in Section 6.4.1, similar numbers of giant planets are recorded as detectable by PLATO, whichever PDF is used to simulate the population. Significantly fewer terrestrial planets are considered detactable by PLATO in simulations based on 'omega' than in simulations based on 'zeta'. This is despite the fact that, as shown in Fig. 6.9, in the bin $0 < log\ R/R_{\oplus} < 0.2$, 'omega' and 'zeta' have similar



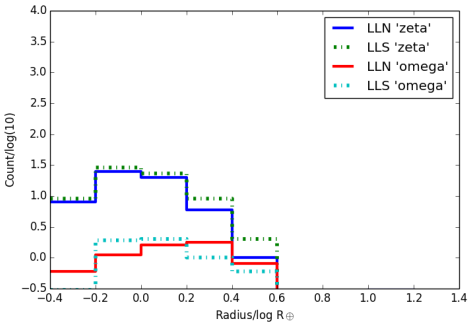
(a) $V < 26$



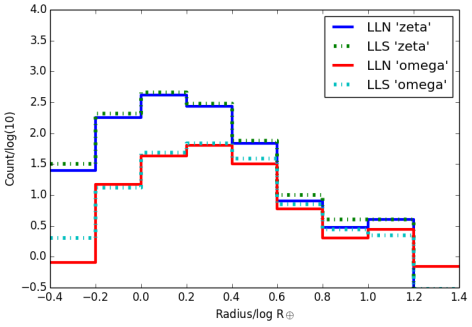
(b) $V < 16$, F5-M



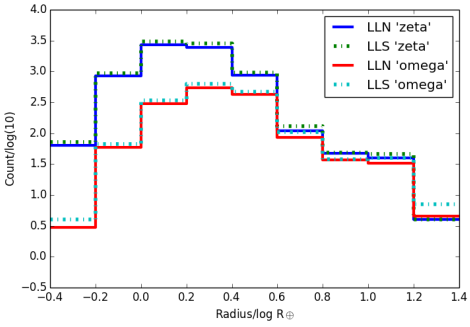
(c) P1



(d) P2

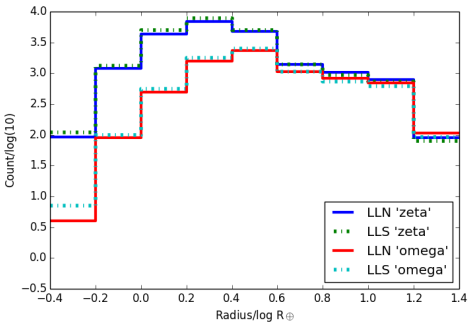


(e) P4

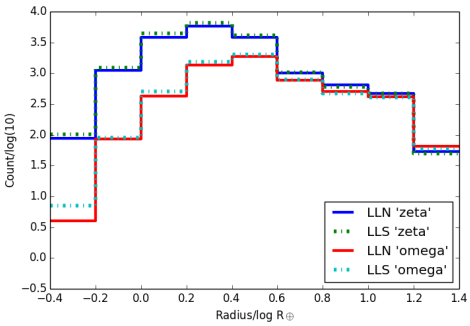


(f) P5

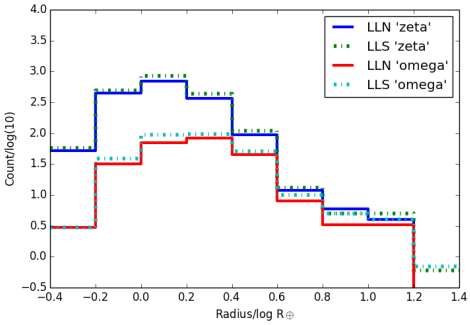
Fig. 6.19 Synthetic planet distribution by radius bin, log bins, 24 cameras, true planet radius. Solid lines: LLN, dash-dot lines: LLS. While ‘zeta’ and ‘omega’ report similar numbers of planets in the giant planet population, in terrestrial planets ‘omega’ reports significantly fewer planets than does ‘zeta’.



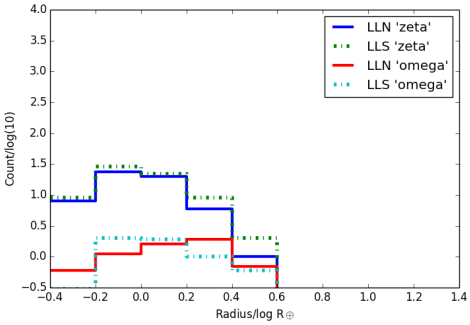
(a) $V < 26$



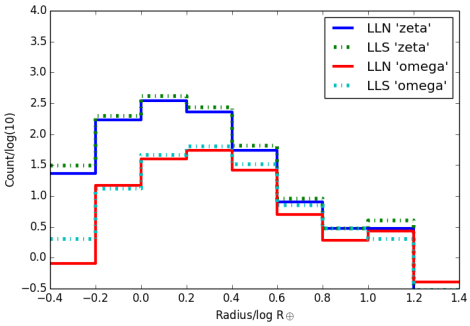
(b) $V < 16$, F5-M



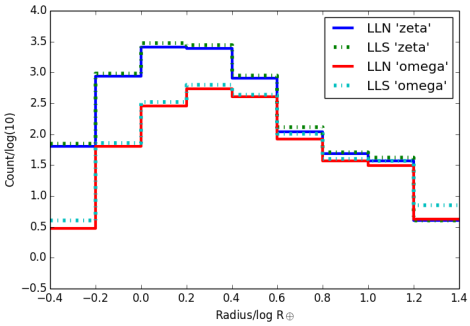
(c) P1



(d) P2

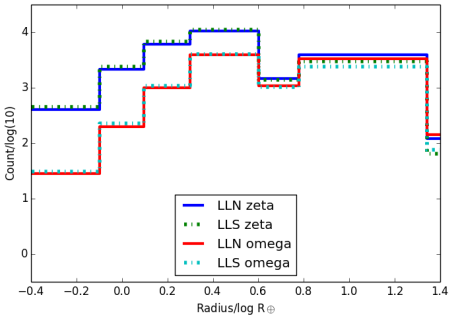


(e) P4

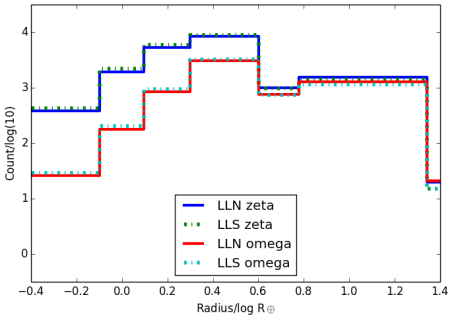


(f) P5

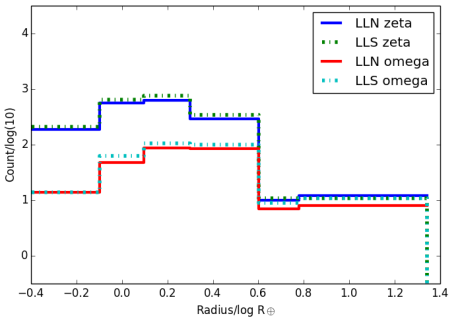
Fig. 6.20 As Fig. 6.19, apparent planet radius after blending.



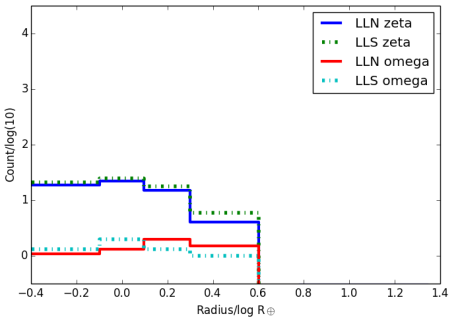
(a) $V < 26$



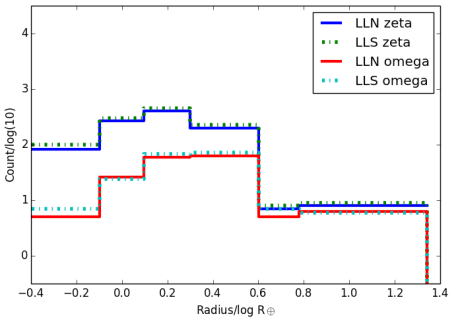
(b) $V < 16$, F5-M



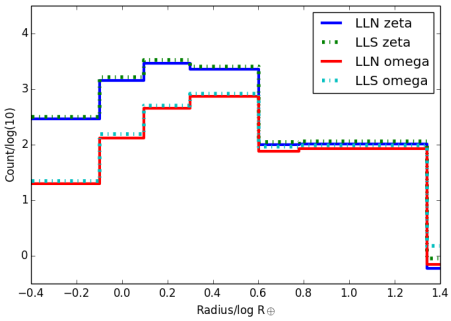
(c) P1



(d) P2

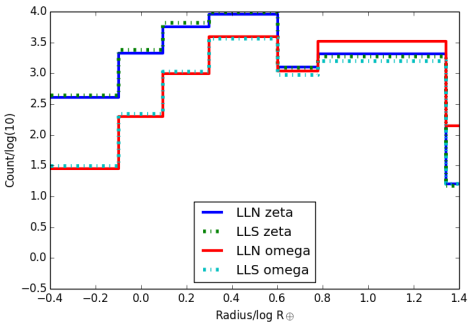


(e) P4

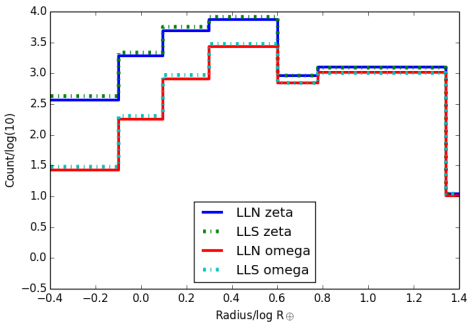


(f) P5

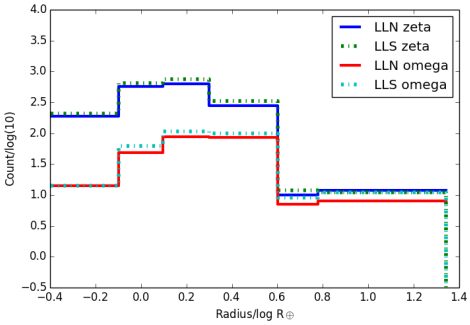
Fig. 6.21 As Fig. 6.19, bins from Fressin et al. (2013), true planet radius.



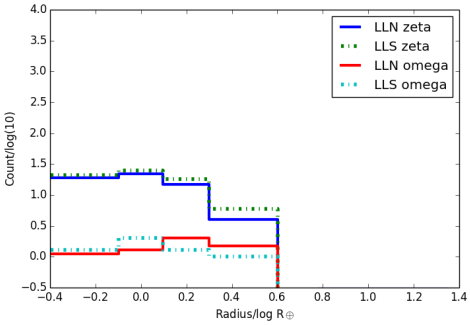
(a) $V < 26$



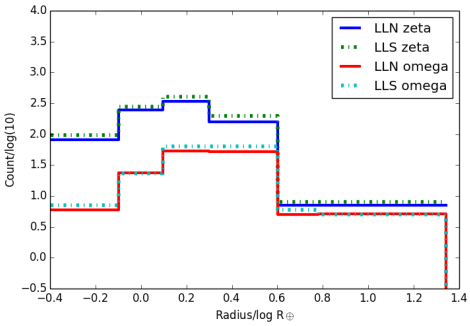
(b) $V < 16$, F5-M



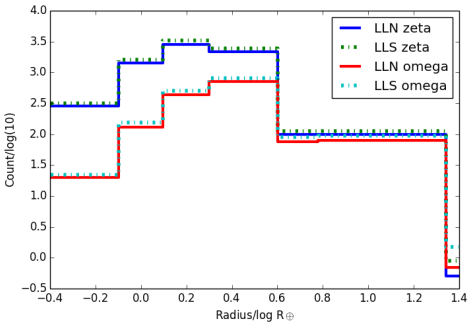
(c) P1



(d) P2



(e) P4



(f) P5

Fig. 6.22 As Fig. 6.21, apparent planet radius after blending.

intrinsic distributions. The difference between the synthetic populations in this bin, an order of magnitude, is the same as the difference between the two normalisation factors.

In each priority population and at most radii, more planets are considered detectable by PLATO in LLS than in LLN (Section 6.4.1). However, in terrestrial planets, the difference between the two fields is more than outweighed by the difference in underlying PDFs.

So there are two fixed points: the similarity in synthetic populations of giant planets when the PDFs are normalised to match the number of observed *Kepler* planets, and the similarity in the intrinsic population (when the PDFs are normalised to 1) in terrestrial planets in the bin $0 < log\ R/R_{\oplus} < 0.2$. For the present, this contradiction is simply presented. Future work will be required to disentangle these elements, and to account for dilution by blending in the *Kepler* field, to approach the true intrinsic exoplanet distribution.

6.5 Summary: Chapter 6

In this chapter, estimated intrinsic exoplanet distributions have been derived, based on the confirmed *Kepler* planet haul. This is done with the caveats that excluding unvalidated candidates may exclude a number of (especially smaller) genuine planets; the effect of dilution by blending with bound and unbound companions to the stellar hosts of confirmed planets is not taken into account in the confirmed *Kepler* planet distribution used in this work; any future work to constrain the radius of stellar hosts may require a re-evaluation of the true radius of the planets; and the completeness of the confirmed *Kepler* planet haul is sensitive both to the completeness of the *Kepler* observations and to the efficiency of the *Kepler* pipeline in identifying shallow transits.

The two distributions derived in this work are believed to bracket the true exoplanet distribution as a function of period and radius. More planets will be observable by PLATO in the Southern field than in the Northern field but, for terrestrial planets, the difference between the two fields is less significant than the difference between the two normalised PDFs.

In the following chapter, the distribution of blended and unblended eclipsing binaries detectable in transiting exoplanet surveys will be estimated: the next step in estimating the ratio of planets to eclipsing binaries detectable by PLATO.

Chapter 7

Eclipsing Binary Distribution

Chapter 1, Section 1.5, discussed the characteristics of eclipsing binaries, in particular those that can be mistaken for transiting exoplanets. This chapter investigates the distribution of binaries as observed in transiting exoplanet surveys, using the *Kepler* main mission as a calibrator for PLATO.

One key requirement for detection is an orbital period less than one third of the total time a field has been observed. The *Kepler* main mission lasted for $\approx$ 1590 days: therefore the effective upper limit is $\approx$ 795 days, assuming a detection at the very beginning and very end of observations and one half way through. Three detections are necessary to prove a possible detection is periodic and not due to the random alignment of two non-periodic events.

After recapping the resources available from the *Kepler* main mission for the study of eclipsing binaries, this chapter will describe the calibration of the synthetic population by mass ratio and by period distribution. Calibrated synthetic populations are then used to extrapolate the likely eclipsing binary haul observable by PLATO.

7.1 Eclipsing binary resources from *Kepler*

Two important data archives/catalogues for the study of eclipsing binaries came out of the *Kepler* main mission: the *Kepler* Eclipsing Binary Catalogue (KEBC) (Prša et al. (2011), Slawson et al. (2011), Matijevič et al. (2012), Conroy et al. (2014b), Conroy et al. (2014a), LaCourse et al. (2015), Kirk et al. (2016) and Abdul-Masih et al. (2016)), and the NASA Exoplanet Science (NExSci) archive[1]. Both these sources have been referenced in previous chapters.

The KEBC takes original *Kepler* light curves and determines if these are unblended eclipsing binaries, in which case they are included in the catalogue and, through the use of the

[1]https://exoplanetarchive.ipac.caltech.edu/

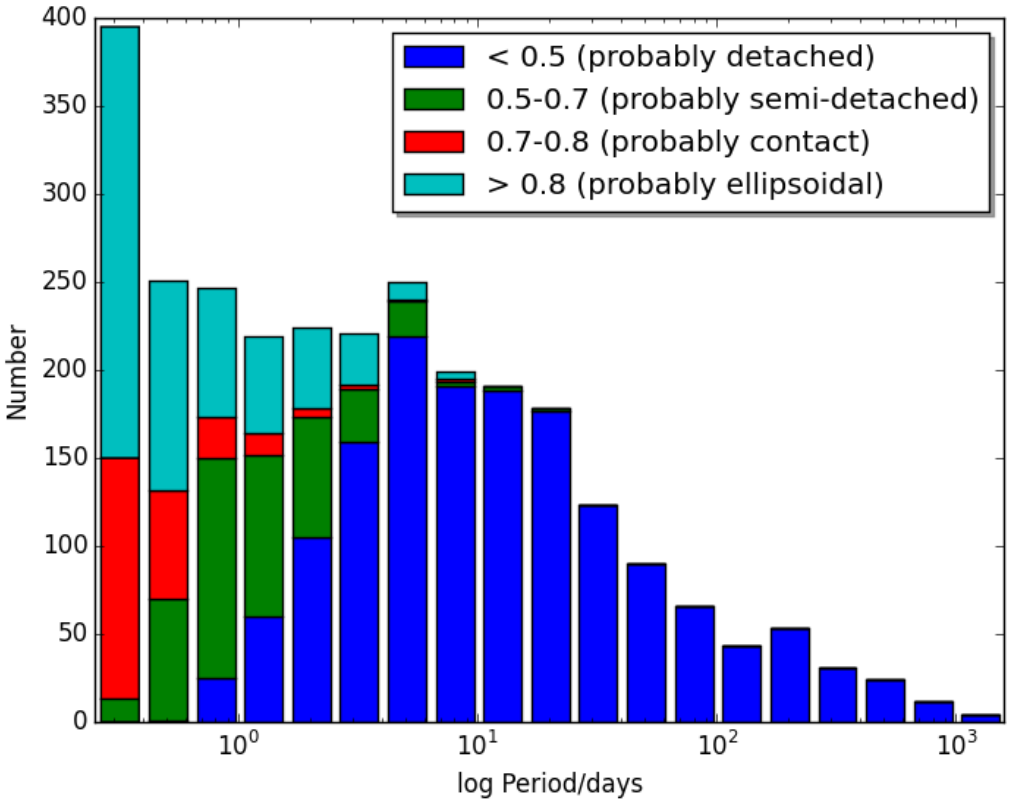


Fig. 7.1 *Kepler* Eclipsing Binary Catalogue, by period and morphology parameter. The morphology parameter is defined in Matijevič et al. (2012). Binaries which are probably detached are most likely to be mistaken for exoplanets. Matijevič et al. (2012) did not claim that the morphology parameter defines precisely what type of binary a given system is, rather that it defines what the system probably is.

morphology parameter (defined in Section 1.5), determines if they are likely to be detached (labelled "probably detached"), semi-detached (labelled "probably semi-detached"), contact (labelled"probably contact") or ellipsoidal binaries (labelled "probably ellipsoidal")[2]. This topic is covered in more detail in Chapter 1.5 and sample light curves are illustrated in Fig 1.7. Data on other objects, such as blended binaries or planets, can also be searched for in the KEBC, although they are not listed in the full catalogue.

Of the 2,876 binaries in the catalogue when the data was downloaded (4 November 2016), 948 (33%) have $P < 1$ day, and a further 1113 (39%) have $1 < P/\text{days} < 10$. This is consistent with the expectation that the shorter the period, the easier an eclipse is to detect.

[2]Ellipsoidal binaries are not a distinct physical class of binaries, but they are unrecognised detached or semi detached binaries, close to Roche lobe filling.

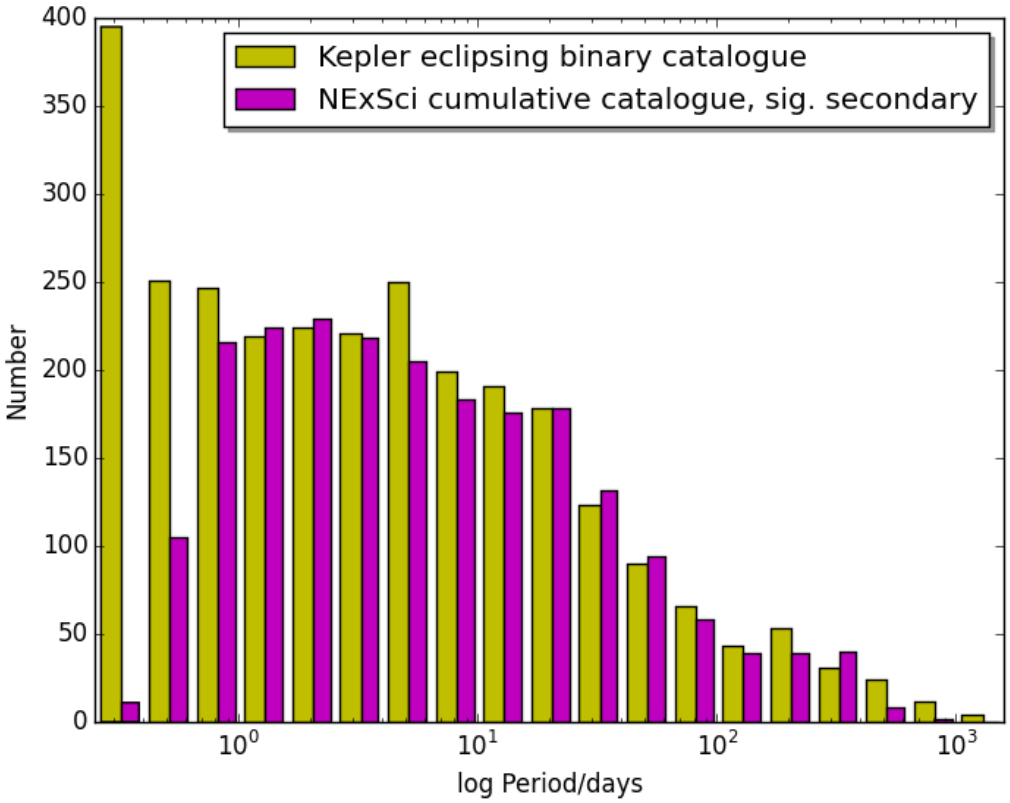


Fig. 7.2 Comparison of number of binaries vs orbital period from the *Kepler* Eclipsing Binary Catalogue (KEBC) and the NASA Exoplanet Science Institute (NExScI). The discrepancy at short periods is in part due to the NExSci cut-off at 0.5 days and in part to the highly sinusoidal nature of many very short period binaries, meaning they are unlikely to be considered as potential planets. KEBC considers all unblended binaries, NExSci only includes those binaries awarded *Kepler* Object of Interest (KOI) status, and may be blended or unblended.

Fig. 7.1 shows the likely breakdown of eclipsing binaries by morphology, remembering of course that ellipsoidal binaries may not be eclipsing. "Probably detached" binaries are most likely to have light curves that could be mistaken for planets, and these tend to have longer orbital periods than other categories. There are very few at < 1 day, while "probably detached" binaries represent about half the total number of binaries in the $0.2 < log\ P/\text{days} < 0.4$ bin (5th bar from the left). "Probably ellipsoidal" binaries dominate at the shortest periods.

Non-eclipsing contact binaries and ellipsoidal binaries appear in catalogues such as the KEBC, which consider all data collected during a transiting exoplanet search, but will be highly sinusoidal and so are unlikely to be considered for follow up as potential planets. Because my simulations only consider eclipsing binaries, non-eclipsing contact and ellipsoidal binaries are excluded by design.

NExSci only lists objects granted KOI (*Kepler* Object of Interest) status, but does list candidates and false positives as well as confirmed planets. This makes it a uniquely useful catalogue. There are two dispositions: "disposition using *Kepler* data", which looks only at the automated results from the *Kepler* pipeline and has two categories, Candidate and False Positive, and "Exoplanet Archive Disposition", which has three categories: Confirmed, Candidate and False Positive. A total of 45 confirmed planets are believed by the *Kepler* pipeline to be false positives in DR25: these cases were manually examined by the *Kepler* False Positive Working Group and in the vast majority of cases were confirmed as planets. Individual objects can be searched for in the *Kepler* Certified False Positive table.[3]

NExSci has four false positive flags:

1. ***NT (not transit like)***. This flag covers alerts arising from threshold crossing events (TCEs) which, on further study of the light curve, were shown not to be transit like. This flag covers stellar variability; starspots; instrumental effects; non periodic detections, such as cosmic ray strikes; and highly sinusoidal light curves, such as those expected from a contact or ellipsoidal binary. Transit timing variations (TTVs) can also trigger this flag.

2. ***SS (significant secondary up to DR24: in DR25 this became stellar eclipse, SE)***. Up to and including DR24 this flag indicated that a significant secondary eclipse had been identified, indicating that this is likely to be an eclipsing binary. From DR25, the definition was changed to indicate that the light curve appeared to indicate a stellar companion rather than a planetary one, whether a secondary eclipse was significant or not. For example, a V-shaped light curve could trigger this flag. Care needs to be exercised, as a V shape by itself is not enough to confirm that the TCE is an eclipsing binary rather than a planet with a high impact parameter.

3. ***CO (centroid offset)***. This flag uses pixel-level diagnostics to determine if the signal is on the target or offset from it. If significantly offset, it is regarded as a false positive on the grounds that it is not on the target star, whether the TCE has been triggered by a blended planet or a blended binary. There is always likely to be a small offset from the calculated difference centroid and the catalogue position of the system being observed. See, for example, Fig. 1.8 in Section 1.5.3, where diagnostics for the confirmed planet *Kepler*-10 c are illustrated. Whether the offset is significant is determined uniquely for each TCE, depending, among other factors, on the number of detections and hence the orbital period. Not all quarters have good quality difference images (see, for example, the discussion of KOI 102.02 in Section 5.2.1), and this will obviously impact the determination of a significant offset. Another factor is the presence of previously unresolvable background stars: if the TCE is on the target but the flux from the previously unknown star is not included in the determination of the centroid offset, then a

[3]https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=fpwg

significant offset can be indicated on the opposite side of the target to the previously unknown star. This can be picked up by the manual vetting carried out by the *Kepler* False Positive Working Group.

4. ***EM (ephemeris match)***. Each TCE is tested by the *Kepler* pipeline to see if there is another entry with matching ephemeris: in particular, is there a match to the orbital period and phase. The separation in terms of rows and columns can be significant, as charge is known to transfer along the same column or the same row of the detector, especially from saturated pixels, resulting in ephemeris matches to distant sources; or it can be much lower, if just the wings of a nearby PSF are detected in the PSF of the target star. The parent may be a variable star or may be an eclipsing binary. The scope of this work does not consider matches to variable stars, such as RR Lyraes: this is a subject for future work. It is also possible for a parent to be flagged as the child of the child, ie two systems cross reference each other as being the parent in an ephemeris match, and manual vetting by the *Kepler* False Positive Working Group determines which TCE is truly the parent and which is the child.

A false positive may carry more than one flag. False positives flagged SS are likely to be eclipsing binaries: those also flagged CO are likely blended eclipsing binaries. With this in mind, the period distribution of entries in the KEBC and NExSci is compared in Fig. 7.2 showing that, apart from short periods where NExSci will be less complete as a period of ≥ 0.5 days is required for further analysis, and very short period binaries often have highly sinusoidal light curves so are unlikely to be considered for follow up as potential planets, the two catalogues are comparable. A false positive flagged CO but not SS could be a blended planet or could be a blended binary in which the secondary eclipse is not observable.

Fig. 7.3 (stellar eclipse) and Fig. 7.4 (centroid offset) show observed *Kepler* false positives from DR25 by derived "planet radius" against period, where the derived "radius" is the radius calculated by the *Kepler* pipeline on the assumption that the TCE refers to an unblended planet and that the 'planet host' is a single star. Both distributions are bimodal, with systems with the stellar eclipse flag showing more systems with a larger derived "radius" (Fig. 7.3), while systems with the centroid offset flag are more common at a smaller derived "radius" (Fig. 7.4). Many systems carrying a stellar eclipse flag have a derived "radius" too large to be considered realistically to be a planet.

There is also a known increase in false positives at around 370 days due to instrumental effects, corresponding to the *Kepler* orbital period, discussed in Section 5.1. This is apparent in the minor long period peak in the upper panels of Fig. 7.4 and Fig. 7.5, and is especially apparent when false alarms (non transiting flag) are also considered.

Confirmation of eclipsing binary status may be through the ephemeris match flag, illustrated in Fig. 7.5. Matching is made possible through the existence of one integrated catalogue from the *Kepler* main mission in which to search for ephemeris matches. Coughlin et al.

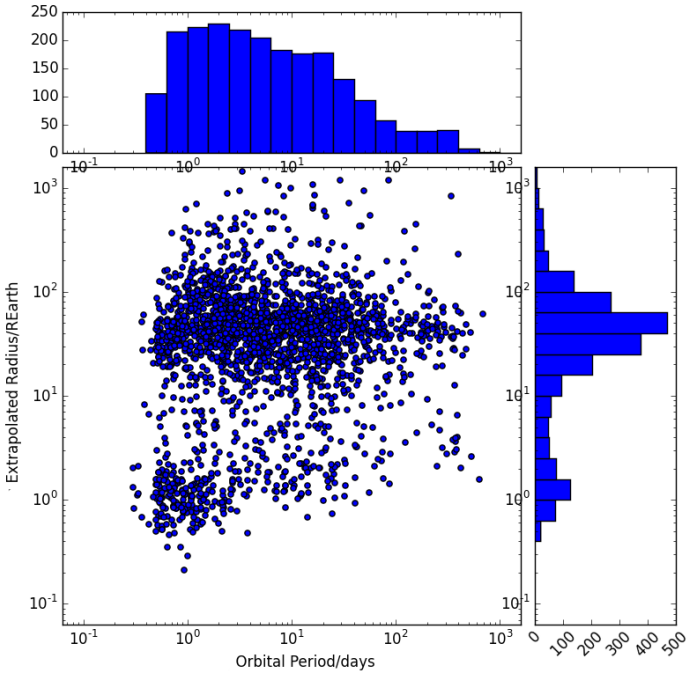


Fig. 7.3 Systems carrying the Stellar Eclipse flag from *Kepler* DR25, extrapolated planet radius against orbital period.

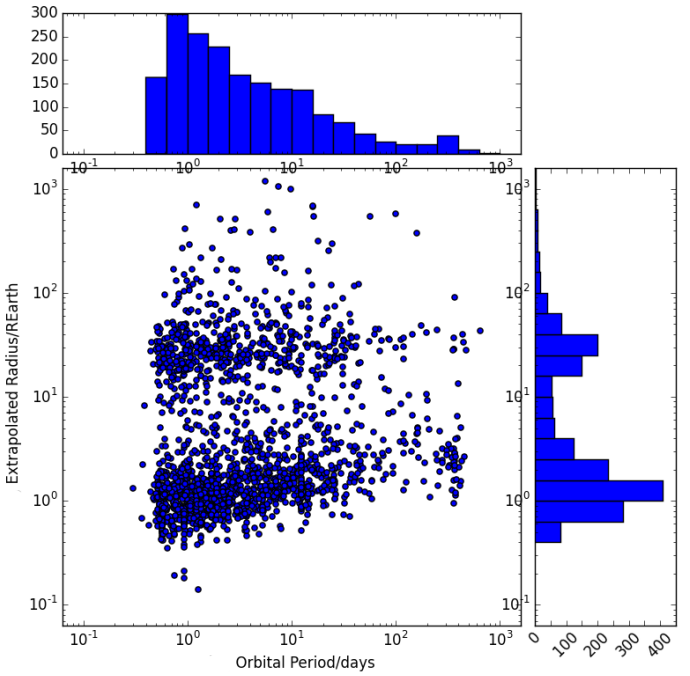


Fig. 7.4 Systems carrying the CO (centroid offset) flag from *Kepler* DR25, extrapolated planet radius against orbital period.

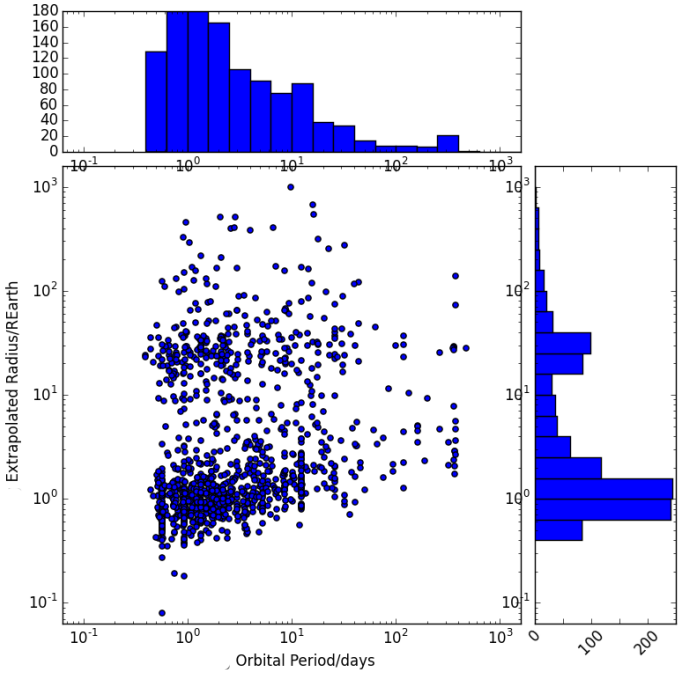


Fig. 7.5 Systems carrying the EC (ephemeris contamination) flag from *Kepler* DR25, extrapolated planet radius against orbital period.

(2014) discuss the analysis of ephemeris matches, and, as shown in Fig. 3 of Coughlin et al. (2014), the parent can be up to 1,000" from the child, with most concentrated between 5"-400" from the child. Ephemeris matches often generate a derived “radius” similar to that of a terrestrial planet because only a small amount of the energy from the parent is seen in the child. “Energy” in this context refers to the incoming photons, converted to photoelectrons in one or more pixels and read out, and includes charge which has transferred along columns or rows due to instrumental effects. The presence of ephemeris matches emphasises the benefits of a comprehensive catalogue of all binaries and variable stars in a field. Gaia DR3 is expected to be able to provide this.

Astrophysical causes of false positives and false alarms which may mimic a terrestrial or Neptune-like planet, in addition to the ephemeris matches to distant systems discussed above, include highly grazing unblended eclipsing binaries (Chapter 1.5.1), unblended eclipsing binaries in which one of the components is a compact object (Chapter 1.5.2), background blends (Chapter 1.5.3), cosmic rays, and cases where only the secondary eclipse is detected. For PLATO, it is extremely unlikely that a cosmic ray will affect all cameras in a group at the same time, therefore this is likely to be less of a factor than it was for *Kepler*. The secondary-only detection scenario is essentially rare. An example of a secondary-only detection is KOI 6064.02, a separate listing of the secondary in a 265.30 day system where the primary has

already been designated KOI 6064.01. Another example is KOI 7968.01, the time series for which is illustrated in Fig. 7.6. In this long period, eccentric binary (P = 424.623 days), there are three detections of the secondary eclipse (indicated by the blue triangles) but only two of the primary eclipse: since three detections are required to trigger further analysis, only the secondary was considered for KOI status. If the second reaction wheel had not failed when it did, or if observations of this system had started in Q3, a third detection of the much deeper primary eclipse could have been expected to be observed.

A few short period planets have also been observed to have secondary eclipses. These include HD 189733 b, TrES-1, HD 209458b and HAT-P-7b (Deming et al. (2006), Burrows et al. (2006), Christiansen et al. (2010)).

7.2 Initial mass ratio calibration

A simple model, assigning random angles of inclination to all eclipsing binaries in a field, and then identifying those that could be 'observed' in a single *Kepler* long cadence integration, did not reproduce the observed distribution by eclipse depth and orbital period as recorded in the KEBC. A single *Kepler* long cadence observation is appropriate when considering the eclipsing binary population observable by *Kepler*, as δF in the KEBC is recorded with a minimum of 0.0001 (100 pm) (see Table 6.1 for the sensitivities of a single long cadence observation). The discrepancy between observations and simulations was particularly noticeable at $P < 10$ days: the observed distribution is approximately flat (Fig.7.1), apart from a peak dominated by "probably ellipsoidal" binaries at the very shortest periods, while the synthetic distribution continued to increase as period decreased.

I sought to match my simulations more closely to observations. This involved weighting by initial mass ratio at $P < 10$ days (Section 7.2.2) and applying a correction factor to the period distribution (Section 7.3). For this, I used the *Kepler* Eclipsing Binary Catalogue. First, however, I examined the population of KOIs in *Kepler* DR24 with the significant secondary flag and $P < 1$ day, to understand the short period boundary, and to further understand what leads to one eclipsing binary being forwarded for further analysis as a possible planet, while another eclipsing binary of similar period was not.

7.2.1 $P < 1$ day

Fig. 7.7 is an annotated plot of false positives with a significant secondary flag, and no other flags, in *Kepler* DR24, downloaded December 2015. This flag implies that the systems in this plot are unblended eclipsing binaries. Very short period giant planets can have secondary eclipses, and where this is the case a great deal can be learned about a planet from the

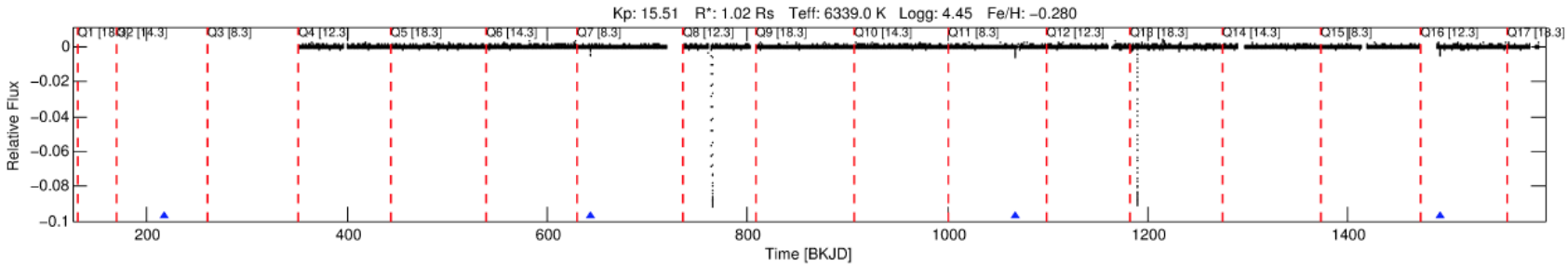


Fig. 7.6 Time series, KOI 7968.01. Blue triangles indicate the eclipses that are being analysed. The triangle in Q2 indicates when an eclipse would have been expected to be observed, if KOI 7968 had been observed at this time (it appears to have only been included in the target list from Q4). Only two primary eclipses were observed, so the primary eclipse does not meet the criteria of three detections required for KOI status. This plot is obtained from publicly available analysis published by NASA (https://exoplanetarchive.ipac.caltech.edu/data/KeplerData/009/009839/009839062/tcert/kplr009839062_q1_q17_dr25_obs_tcert.pdf, top of page 2 (DV One Page Summary)).

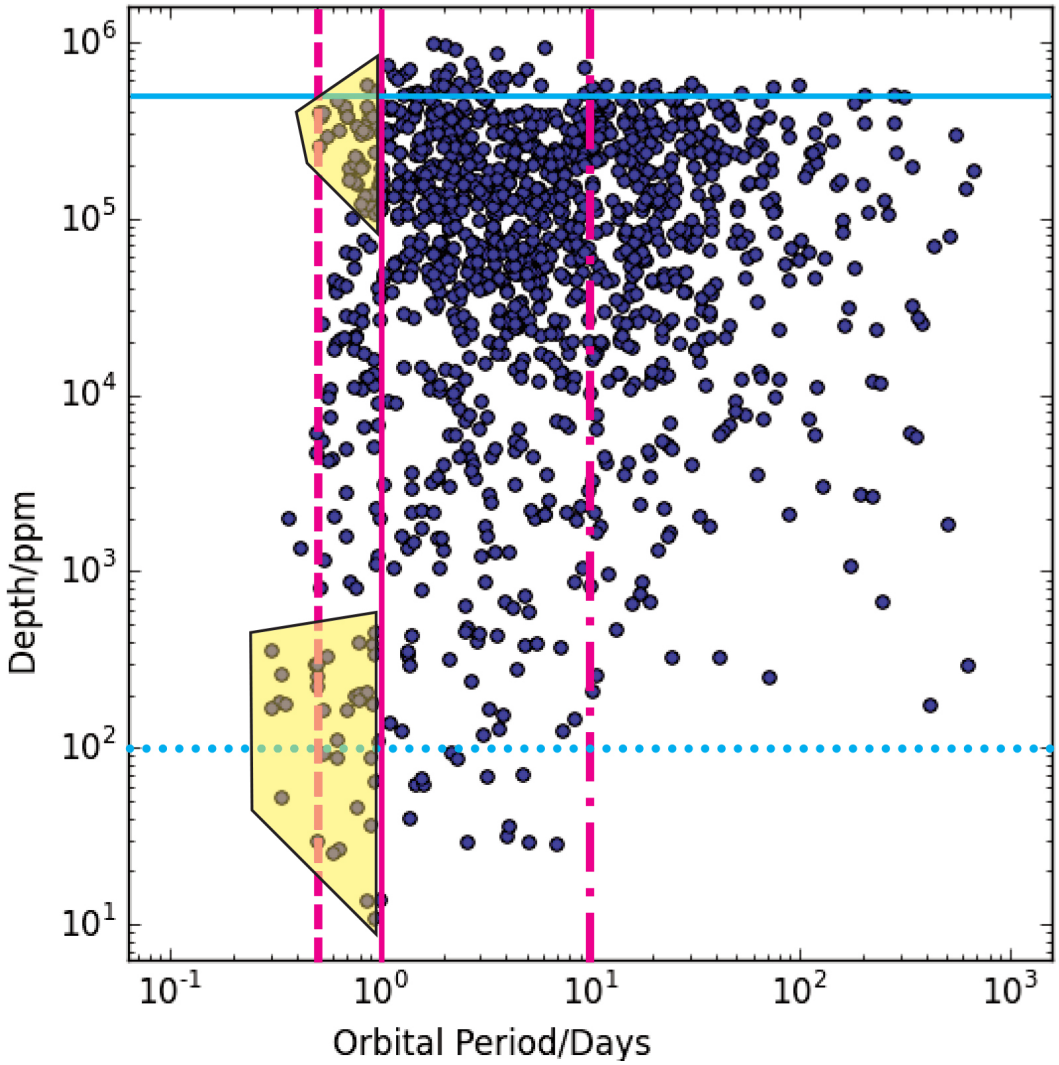


Fig. 7.7 False positives with the Significant Secondary flag and no other false positive flags, *Kepler* DR24, transit depth against period. $P = 0.5$ days: dashed magenta line. This is the minimum period at which KOIs were routinely examined. $P = 1$ day: solid magenta line. $P = 10$ days: dash-dot magenta line. Eclipse depth = 500,000 ppm: solid cyan line. This is the the "theoretical" maximum eclipse depth. Eclipse depth = 100 ppm: dotted cyan line. This is the minimum eclipse depth recorded in the *Kepler* Eclipsing Binary Catalogue. The regions indicated in yellow at $P < 1$ day are dominated by $b < 1$, the other regions at $P < 1$ day are dominated by $b > 1$ (grazing eclipses). Note that binaries with $b > 1$ can be found in the regions identified in yellow, and binaries with $b < 1$ can be found in the region not marked in yellow.

secondary eclipse (Haswell, 2010). The significant secondary flag was only applied to those sources believed not to be planets. Note that, according to the KEBC, the true period of a proportion of these binaries is double that stated on NExSci. DR24 is more suitable for this work than DR25, because of the change in definition of the SS flag from Significant Secondary to Stellar Eclipse: the latter includes V-shaped light curves without an observable secondary eclipse, so may include grazing eclipses of giant planets, which are also V-shaped.

Three periods are marked in Fig. 7.7. P = 10 days (dash-dot line) is marked as periods shorter than this can generally be expected to have circularised or be in the process of circularising, depending on the age of the system. These are the most suitable binaries for comparison with a `BiSEPS` simulation, which utilises zero eccentricity. P= 0.5 days (dash line) is the minimum period at which sources were routinely examined by the *Kepler* pipeline. Sources for which shorter periods have been recorded have had their true periods manually identified. P = 1 day (solid line) marks the upper limit by period of those binaries investigated to understand the nature of very short period binaries which may be identified as worthy of further investigation as potential planets in transiting exoplanet surveys. Fig. 7.2 shows that the greatest discrepancy between the *Kepler* Eclipsing Binary Catalogue and false positives with significant secondaries/stellar eclipses listed on NExSci is at $P < 1$ day.

An investigation of false positives with the significant secondary flag at $P < 1$ day revealed one particularly interesting characteristic, the impact parameter b. b is derived from examination of the light curve and is < 1 when the centre of the occluding star passes over the disc of the occluded star at conjunction and is > 1 when the centre of the occluding star does not pass over the disc of the occluded star, according to the definitions used by NExSci. b may change from one data release to another, and this analysis relates specifically to DR24. Regions in the *log* depth *log P* parameter space with $P < 1$ day where sources with $b < 1$ were clustered are indicated in Fig. 7.7 by yellow shading. Note that some systems with $b > 1$ fall in these yellow regions and some systems with $b < 1$ fall in the unshaded region. There is overlap. This does, however, indicate two different types of eclipsing binary with $b < 1$: the shallower eclipses with $b < 1$ may be due to eclipsing binaries in which one component is a compact object, or to background eclipsing binaries where the offset is not flagged as significant.

At the shortest periods, there is an apparent boundary in the population with $b > 1$, which increases in depth with increasing period. Tests indicated this could be reproduced in a synthetic population by restricting the transit duration T_{dur} as a fraction of orbital period P, $\delta T = T_{dur}/P$, to < 9%, and considering grazing binaries only. Analysts at NASA had found that $\delta T \lessapprox 10\%$ was required for such a transit to be detectable (J. Christiansen, staff scientist at the NASA Exoplanet Archive, priv. comm), which is in line with our findings. There are cases with longer δT listed on NExSci: these are often cases where the larger body is believed to be an evolved star, such as a red giant, and may not be grazing.

Two eclipse depths are highlighted on Fig. 7.7. The lower one, 100 ppm (cyan dotted line), is the minimum depth recorded in the KEBC. The KEBC was the catalogue used in this study to calibrate by initial mass ratio distribution and by period distribution. NExSci includes shallower transits, as here the aim was to identify potential planets: the Earth transiting the Sun would have a transit depth of 84 ppm.

The other eclipse depth indicated in Fig. 7.7, 500,000 ppm (cyan solid line), is the "theoretical" upper limit to eclipse depth, based on two equal radius stars at an angle of inclination of 90%: the maximum eclipse depth is 50% of the out of transit flux when one star completely occludes the other. In fact, as can be seen from Fig. 7.7, some stars have been observed to exceed this limit. Examining a synthetic population of eclipsing binaries, two groups of systems that could exceed this limit were identified.

In the first group, a maximum eclipse depth of 60% could be found in systems which had not experienced mass transfer, but in which one star was significantly larger than the other, often in a MS/MS pair, and in which the system was viewed edge on, or close to edge on. In such an eclipse, the occluding star passes over the brightest part of the occluded star and so, due to the effects of limb darkening, the depth of the eclipse can exceed 50%. Limb darkening is described in more detail in Section 3.1.1.

In the second group, a maximum eclipse depth of 85% could be found in synthetic systems which experienced mass transfer through Roche lobe overflow between two stars in an MS/MS, RGB/MS or RGB/RGB pairing. In Fig. 7.7 the maximum recorded eclipse depth is of a binary with a period of 1.81 days (KOI 5614.01/KIC 9101279, depth = 961,488 $\pm$2346 ppm, in *Kepler* DR24, depth revised in *Kepler* DR25 to 921671 $\pm$281 ppm). There appears to be an increase in eclipse depth in the observed population with orbital period from $P = 0.5$ days to $P = 1.8$ days. Matter transferring from a donor star to an accretor will form a third source of light, especially if an accretion disc with a bright spot is formed. If the bright spot is obscured during an eclipse, the change in flux will be greater than would be expected when considering two stars alone.

7.2.2 $P < 10$ days

Using the KEBC, the role of the initial mass ratio in producing the observed distribution was explored in systems with $P < 10$ days.

The detection probability υ

The initial mass ratio distribution ε is calculated in BiSEPS as

$$\varepsilon = (s+1)q^s \tag{7.1}$$

where q is the mass ratio and s is a free parameter specified as an input for each population simulation.

The probability of detecting an eclipsing binary depends on the transit probability and observed brightness of the system, and hence can be estimated as

$$\upsilon \propto \varepsilon \left(\frac{R_1 + R_2}{a} \right) \left(\frac{L_1 + L_2}{d^2} \right) \tag{7.2}$$

where R_1 and R_2 are the radii of the two stars and a is the semi major axis, L_1 and L_2 are the luminosity of the two stars (measured in $L_\odot$) and d is the distance from Earth in parsecs.

The usefulness of Eq. 7.2 as an estimate for detection probability was tested as follows. Using data from the BiSEPS model archive (Chapter 2), the parameters of the secondary were determined for a detached system in which the primary was a representative star (1 $R_\odot$, 1 $M_\odot$, 1 $L_\odot$, age = 4.6 Ga), for secondaries of the following initial mass: 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0 $M_\odot$, all at age = 4.6 Ga. The distance d and separation a were treated as constants. The corresponding value of υ for each of these data points was calculated for $s = 0$, and then normalised to 1. Using the synthetic *Kepler* field from Farmer et al. (2013), a synthetic eclipsing binary population was obtained. The distribution by mass ratio was obtained and normalised for $0.1 < q < 1$, always taking the evolutionary secondary over the evolutionary primary to derive q. All binaries in which the evolutionary primary was a compact remnant were removed from analysis, as the present q will reflect mass lost from the system: these represent $\approx 0.4\%$ of systems in an average simulation interpreted with Kepler single long cadence precision on the field from Farmer et al. (2013), where $s = 0$. In addition, systems in which neither component was a compact remnant but which, as a result of mass transfer, had $q > 1$, were also removed from the analysis, as they are not representative of the systems used to derive υ for this comparison. In these systems, the evolutionary secondary is now the more massive star (for example, Algol-like systems). These represent $\approx$ 3% of eclipsing binary systems in an average simulation interpreted with *Kepler* single long cadence precision on the field from Farmer et al. (2013), where $s = 0$. The value of $s = 0$ was used in this comparison, as this matches the parameter with which the field was created. The results are presented and demonstrate satisfactory agreement in Fig. 7.8.

To generalise this comparison I explored the effect of different primary mass, and different evolutionary age of the primary, from ZAMS to TAMS on υ.

The normalised υ was calculated at ZAMS and at TAMS for stars with a primary mass of $0.5 < M/M_\odot < 2.0$ at intervals of 0.1 $M_\odot$. The results for $s = 0$ are given in Table 7.1 and Table 7.2. Fig. 7.9 and Fig. 7.10 compare υ for the same range of properties of the primary as in Table 7.1 and Table 7.2 for three values of s, -0.5, 0 and 0.5. There is noticeably more spread at TAMS than at ZAMS, partly because main sequence lifetime decreases as initial

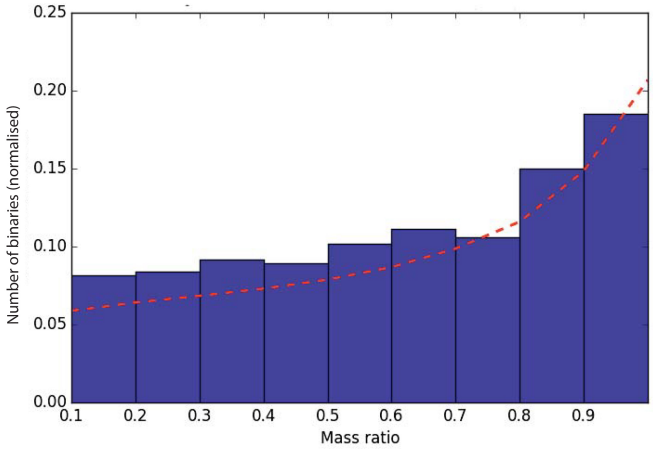


Fig. 7.8 Normalised detection probability υ (Eq. 7.2) derived with respect to $s = 0$ (red line) compared with an eclipsing binary simulation in a synthetic field from Farmer et al. (2013), interpreted with *Kepler* precision in a 30 minute integration, over mass ratio (see text for more details). The properties of the primary in υ are fixed in this comparison at 1 $R_\odot$, 1 $M_\odot$, 1 $L_\odot$, age = 4.6 Ga. The properties of the secondary are based on stars of the following initial mass: 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0 $M_\odot$, all at age = 4.6 Ga. Systems simulated but excluded from this analysis are those in which one component is a compact remnant ($\approx$ 0.4%) and systems without a compact component but with $q > 1$ ($\approx$ 3%).

Table 7.1 Normalised detection probabilities, $s = 0$. Horizontal: mass ratio. Vertical: initial mass of the primary, $M_\odot$. Time point: Zero Age Main Sequence (ZAMS).

Primary initial mass	**q =0.1**	**q = 0.2**	**q = 0.3**	**q = 0.4**	**q = 0.5**	**q = 0.6**	**q = 0.7**	**q = 0.8**	**q = 0.9**	**q = 1.0**
0.5	0.049	0.058	0.067	0.076	0.086	0.098	0.112	0.128	0.149	0.176
0.6	0.051	0.060	0.068	0.076	0.085	0.095	0.107	0.123	0.149	0.185
0.7	0.053	0.062	0.069	0.075	0.082	0.092	0.104	0.122	0.149	0.192
0.8	0.054	0.063	0.068	0.074	0.081	0.089	0.102	0.122	0.151	0.195
0.9	0.055	0.063	0.068	0.073	0.080	0.089	0.102	0.122	0.151	0.196
1.0	0.057	0.063	0.068	0.073	0.079	0.089	0.102	0.121	0.150	0.198
1.1	0.057	0.063	0.067	0.072	0.079	0.088	0.101	0.120	0.151	0.201
1.2	0.058	0.063	0.067	0.072	0.079	0.087	0.100	0.120	0.153	0.203
1.3	0.058	0.062	0.066	0.072	0.079	0.087	0.099	0.119	0.153	0.205
1.4	0.058	0.062	0.066	0.071	0.078	0.086	0.099	0.120	0.155	0.204
1.5	0.057	0.061	0.065	0.071	0.077	0.085	0.100	0.123	0.159	0.201
1.6	0.056	0.061	0.065	0.071	0.077	0.086	0.100	0.127	0.158	0.199
1.7	0.056	0.060	0.065	0.071	0.077	0.086	0.104	0.128	0.157	0.196
1.8	0.055	0.060	0.065	0.071	0.077	0.088	0.105	0.128	0.157	0.195
1.9	0.055	0.059	0.065	0.071	0.077	0.089	0.107	0.129	0.156	0.192
2.0	0.055	0.059	0.065	0.070	0.077	0.090	0.108	0.128	0.156	0.191

Table 7.2 Normalised detection probabilities, $s = 0$. Horizontal: mass ratio. Vertical: initial mass of the primary, $M_\odot$. Time point: Terminal Age Main Sequence (TAMS).

Primary initial mass	**q =0.1**	**q = 0.2**	**q = 0.3**	**q = 0.4**	**q = 0.5**	**q = 0.6**	**q = 0.7**	**q = 0.8**	**q = 0.9**	**q = 1.0**
0.5	0.049	0.059	0.067	0.076	0.086	0.098	0.111	0.128	0.148	0.178
0.6	0.052	0.061	0.069	0.076	0.084	0.094	0.106	0.121	0.147	0.189
0.7	0.055	0.064	0.070	0.076	0.082	0.091	0.101	0.119	0.143	0.200
0.8	0.057	0.065	0.070	0.075	0.081	0.087	0.099	0.115	0.143	0.208
0.9	0.061	0.068	0.071	0.075	0.079	0.085	0.093	0.106	0.133	0.228
1.0	0.064	0.068	0.071	0.074	0.078	0.083	0.090	0.103	0.132	0.236
1.1	0.064	0.068	0.070	0.073	0.077	0.083	0.090	0.102	0.135	0.237
1.2	0.064	0.067	0.070	0.073	0.077	0.083	0.090	0.104	0.133	0.238
1.3	0.063	0.066	0.069	0.072	0.077	0.082	0.090	0.105	0.140	0.236
1.4	0.063	0.066	0.069	0.072	0.077	0.081	0.090	0.106	0.141	0.235
1.5	0.061	0.064	0.067	0.070	0.075	0.080	0.091	0.114	0.149	0.228
1.6	0.063	0.066	0.068	0.072	0.075	0.081	0.091	0.109	0.138	0.236
1.7	0.063	0.066	0.068	0.072	0.075	0.081	0.092	0.109	0.136	0.237
1.8	0.063	0.065	0.068	0.072	0.075	0.081	0.092	0.109	0.139	0.236
1.9	0.061	0.064	0.066	0.070	0.073	0.081	0.094	0.114	0.147	0.231
2.0	0.063	0.065	0.068	0.071	0.074	0.082	0.093	0.108	0.140	0.236

stellar mass increases, but s is clearly more significant than the initial mass of the primary. Fig. 7.11 and Fig. 7.12 extend this analysis to higher values of s, again with more spread at TAMS than at ZAMS, especially at higher mass ratios.

This indicates that the initial mass ratio distribution, characterised by s, plays a crucial role in identifying properties of an eclipsing binary population in transiting exoplanet surveys that target Sun-like stars as, when normalised, the detection probability is largely independent of the properties of the primary.

Applying the detection probability υ to the observed sample

Using the detection probability υ, the observed population from the *Kepler* Eclipsing Binary Catalogue was examined to determine the most likely exponent to the initial mass ratio distribution, s, in unblended eclipsing binaries with $P < 10$ days.

The restriction to $P < 10$ days arose because BiSEPS is limited to zero eccentricity, and most binaries with $P < 10$ days are circularised. This is not always the case: at the time of writing (22 March 2018) there are 41 certified false positives listed on NExSci with "Heartbeat" or "Heartbeat?" in the comment and nine of these have periods of less than 10

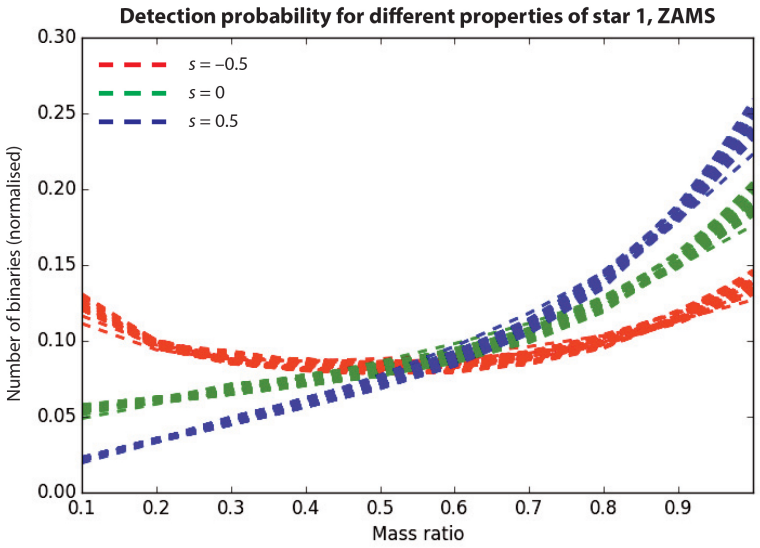


Fig. 7.9 Normalised detection probability, $s = -0.5$, $s = 0$ and $s = 0.5$, ZAMS. Primary: $0.5 < M/\mathrm{M}_{\odot} < 2.0$, $0.1 < q < 1.0$ (similar to Table 7.1). s is more significant than the properties of the primary.

days.[4] A heartbeat star is highly eccentric and the primary and secondary eclipses occur in close proximity to one another: the resulting light curve resembles an ECG signal, hence the name "heartbeat" star (Shporer et al. (2016), Hambleton et al. (2016), Fuller (2017)). One factor is age: younger stars are likely to be in the process of circularising rather than already circularised. However, for the bulk of the population, the assumption of low or zero eccentricity for $P < 10$ days is generally correct.

For $0.45 < P/\mathrm{days} < 10.1$, 971 binaries were identified within the KEBC with the following data: a primary eclipse depth, a secondary eclipse depth, and an estimate of T_{eff} of the primary. A field of 1 deg^2 was selected from the full Galaxy simulation discussed in Chapter 4, centred on $l = 50°$, $b = 8°$, $Kp \leq 26$, and this area was searched for models which matched the parameters of the 971 systems from the KEBC. The allowable variation was 5% in effective temperature and 10% for period, primary eclipse depth and secondary eclipse depth.

For each of the 971 binaries, all matching models within the selected square degree were identified. The proportion for each of the matching models that would fall within a bin of width 0.1 in $0.1 < q < 1.0$ was identified. The resulting distribution over mass ratio indicated a high value of s, possibly as high as $s = 5$.

To test this finding, a series of additional synthetic fields were generated in the same region of the sky, with a limiting magnitude of $Kp \leq 16$, with the following values of s: -0.5,

[4] https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=fpwg

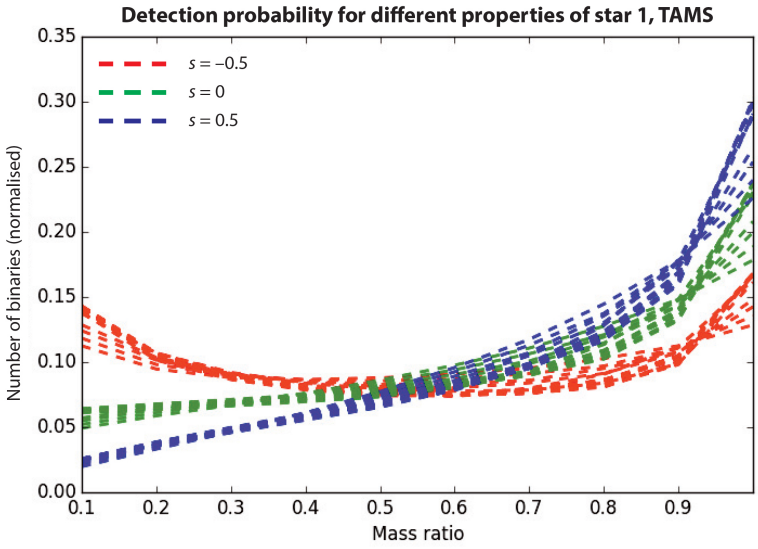


Fig. 7.10 Normalised detection probability, $s = -0.5$, $s = 0$ and $s = 0.5$, TAMS. Primary: $0.5 < M/\mathrm{M}_{\odot} < 2.0$, $0.1 < q < 1.0$ (similar to Table 7.2). s is more significant than the properties of the primary, but the spread within s is more significant at TAMS than at ZAMS.

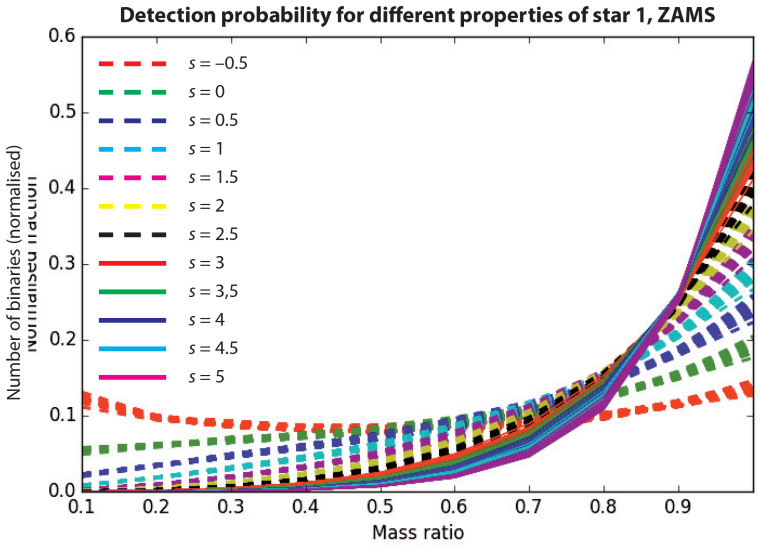


Fig. 7.11 Normalised detection probability, $s = -0.5$ to $s = 5.0$, ZAMS. Primary: $0.5 < M/\mathrm{M}_{\odot} < 2.0$, $0.1 < q < 1.0$ (similar to Table 7.1). s is more significant than the properties of the primary.

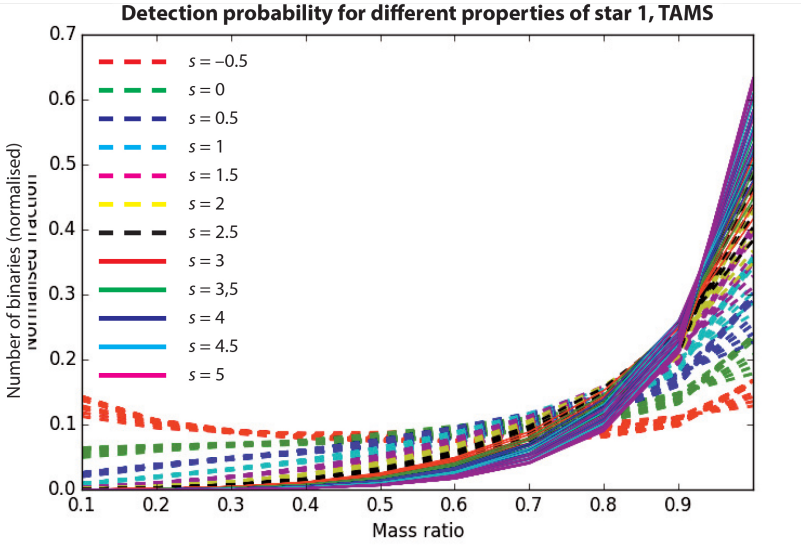
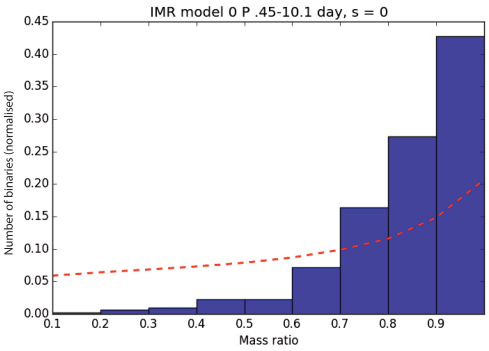


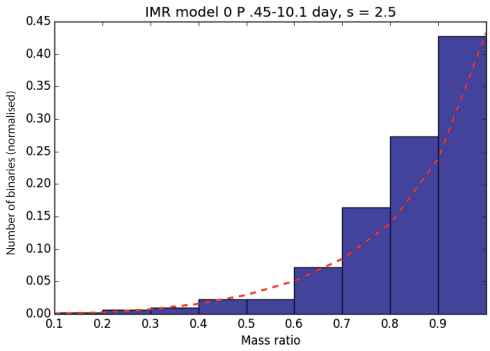
Fig. 7.12 Normalised detection probability, $s = -0.5$ to $s = 5.0$, TAMS. Primary: $0.5 < M/\mathrm{M}_{\odot} < 2.0$, $0.1 < q < 1.0$ (similar to Table 7.2). s is more significant than the properties of the primary. s is more significant than the properties of the primary, but the spread within s is more significant at TAMS than at ZAMS.

0., 0.5, 1., 1.5, 2., 2.5, 3., 3.5, 4. and 4.5. The model matching process was carried out for each of these fields, and the resulting distribution over mass ratio for each of these fields was compared with the υ derived from $-0.5 < s < 4.5$, age = 4.6 Ga, properties of the primary 1 $\mathrm{M}_{\odot}$, 1 $\mathrm{R}_{\odot}$, 1 $\mathrm{L}_{\odot}$. Two values were sought: the minimum acceptable match on an $s = 0$ base, as used in the Galaxy simulations described in Chapter 4, and the most likely true distribution. This was assessed both from visual inspection of the resulting plots (Fig. 7.13, for example) and a Kolmogorov-Smirnov two-sample test, where a low K-S statistic and high p-value indicates the best match (Table 7.3). The conclusions from the K-S test are weak.

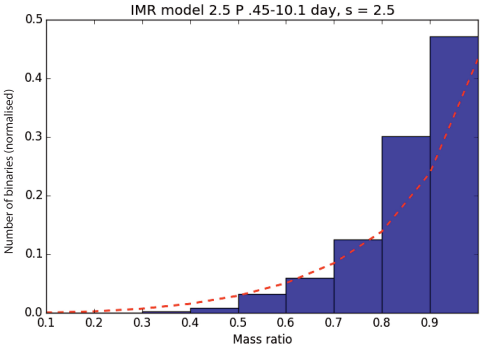
Fig. 7.13, panel (a), shows an $s = 0$ υ curve on an $s = 0$ population base and it is clear that the resulting distribution over mass ratio in no way matches $\upsilon(q)$. This is reflected in the low p-value for the same parameters in Table 7.3, where a high p-value indicates that the hypothesis that the two samples are drawn from the same distribution cannot be rejected. From the K-S two sample test, the best match for a base with $s = 0$ is υ calculated with $s = 2$, slightly better than for $s = 2.5$ and $s = 3$ (Table 7.3). However, visual inspection of the bar charts indicates that a base of $s = 0$ with υ derived using $s = 2$ is not a good match in the most populated bin, $0.9 < q < 1$. The curve for $s = 2.5$ on the same base, $s = 0$, is a much better visual fit. It is illustrated in Fig. 7.13, panel (b). It also produces a reasonable match in the K-S two sample test.



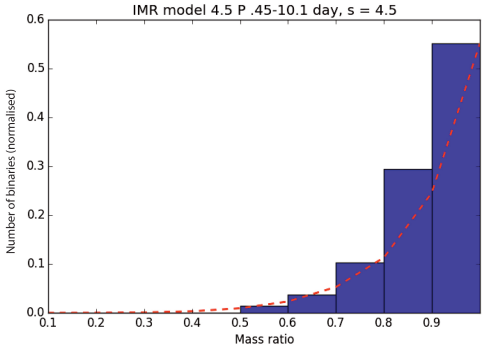
(a) Base: $s = 0$. υ: $s = 0$



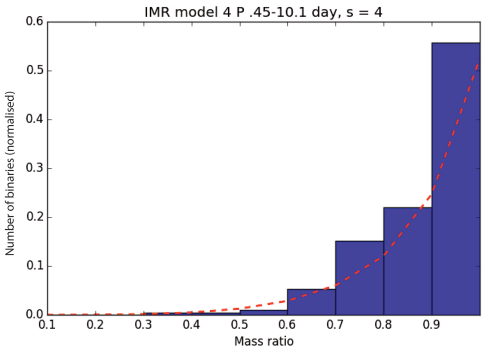
(b) Base: $s = 0$. υ: $s = 2.5$



(c) Base: $s = 2.5$ υ: $s = 2.5$



(d) Base: $s = 4.5$ υ: $s = 4.5$



(e) Base: $s = 4$. υ: $s = 4$.

Fig. 7.13 Synthetic populations matched to data from the Kepler Eclipsing Binary Catalogue, to derive distribution over mass ratio. The synthetic fields were generated with different values of s, as indicated in the panel label ('base') along with the value of s used to derive the normalised detection probability υ (red line). υ from $s = 2.5$ is the minimum acceptable fit to a field simulated with an $s = 0$ base (panel (b)), and also is an acceptable, but not perfect, fit on a field simulated with an $s = 2.5$ base (panel (c)). The highest value of s used was 4.5, and this is illustrated in panel (d): the fit here is better. This suggests that in short period binaries, there is a trend towards equal mass components. The Kolmogorov-Smirnov two sample test (Table 7.3) indicates that $s = 2.5$ (panel (c)) and $s = 4.0$ (panel (e)) represent the best fit, but $s = 4.5$ is the only one with a good match at $0.9 < q < 1$ (panel (d)).

Where υ is calculated with the same value of s as was used to generate the synthetic field, the best result in the K-S two sample test is achieved when $s = 4$, closely followed by $s = 2.0$ and $s = 2.5$, although the K-S statistic in these cases is not particularly good. However, when the plots are inspected, there is again a poor match in the most populated bin, $0.9 < q < 1.0$. To obtain a good match in this bin on visual inspection, it is necessary to go to $s = 4.5$: however, the K-S two sample test indicates that this is a poorer match overall. A contributory factor to this poor result is that four period bins from $0.1 < q < 0.2$ to $0.4 < q < 0.5$ are empty in the simulated population when the field was simulated with $s = 4.5$. The suitability of the K-S test in this case is questionable, when the K-S statistics contradict the visual inspections.

Short period binaries are believed to be an atypical group. Systems cannot easily form at $a < 1$ AU (Moe and Di Stefano (2017) and references therein). For equal mass binaries, this indicates a minimum period at formation of ≈ 260 days where both stars are 1 $M_\odot$, ≈ 370 days where both stars are 0.5 $M_\odot$ and ≈ 180 days where both stars are 2 $M_\odot$, so, at $P < 10$ days, migration must have occurred. Raghavan et al. (2010) demonstrated that the inner pairs of triple systems are most likely to be equal mass, while the component masses in the general binary population are more evenly distributed. For this reason, although $s = 4$ offered the best match to the observed population with $P < 10$ days, the $s = 0$ base is used for the full PLATO simulations, as it appears to be more widely applicable, with short period binaries weighted to match the results of this experiment on the $s = 0$ base. This weighting was achieved by binning the synthetic eclipsing binary population by q, and weighting each bin by $s(2.5, q)/s(0.0, q)$, where $s(2.5, q)$ and $s(0.0, q)$ are based on the model with a Sun-like primary (1 $R_\oplus$, 1 $M_\oplus$, 1 $L_\oplus$, age = 4.6 Ga). Given the similarity between normalised υ shown in Fig. 7.11 and Fig. 7.12 at different s, the model was selected arbitrarily. There are, however, advantages in taking a model approximately half way through its main sequence life, as there is more spread at TAMS than at ZAMS.

This weighting is done as a final step: as this work continues, the aim is to apply a weighting at the beginning of the simulations.

Having weighted the short period binaries by initial mass ratio distribution, the next step was to consider calibration by period for all binaries.

7.3 Period distribution calibration

While it is important to regard the *Kepler* Eclipsing Binary Catalogue as a calibrator, it is also important to remember that this catalogue is not complete. Not all stars in the field were observed, and not all binaries in the observed part of the field made it into the catalogue, as the signal needed to be detectable by *Kepler*. In the case of binaries, integration time is not as significant as it was for planets: most binaries in the KEBC fall within the sensitivity for a

Table 7.3 Results of a Kolmogorov-Smirnov 2 sample test on the calibration of the mass ratio distribution of systems with $P < 10$ days. A low K-S statistic and a high p-value means it is not possible to reject the hypothesis that the distributions of the two samples are the same. In this case, the two samples are the resulting distribution over mass ratio, and the normalised detection probability, υ. "Base" refers to the value of s when each field was simulated. The results from this K-S test are weak.

Base	υ	K-S	p	Base	υ	K-S	p	Base	υ	K-S	p
-0.5	-0.5	0.67	0.015	0	-0.5	0.67	0.015	0.5	-0.5	0.67	0.015
-0.5	0	0.44	0.23	0	0	0.56	0.067	0.5	0	0.56	0.067
-0.5	0.5	0.44	0.23	0	0.5	0.56	0.067	0.5	0.5	0.44	0.23
-0.5	1.	0.34	0.53	0	1.	0.36	0.49	0.5	1.	0.34	0.53
-0.5	1.5	0.24	0.89	0	1.5	0.26	0.86	0.5	1.5	0.24	0.89
-0.5	2.	0.22	0.95	0	2.	0.16	1.0	0.5	2.	0.14	1.0
-0.5	2.5	0.22	0.95	0	2.5	0.20	0.98	0.5	2.5	0.14	1.0
-0.5	3	0.22	0.95	0	3	0.20	0.98	0.5	3.	0.12	1.0
-0.5	3.5	0.26	0.87	0	3.5	0.30	0.70	0.5	3.5	0.16	1.0
-0.5	4	0.28	0.79	0	4	0.30	0.70	0.5	4.	0.19	0.99
-0.5	4.5	0.28	0.79	0	4.5	0.30	0.70	0.5	4.5	0.19	0.99
1.	-0.5	0.67	0.015	1.5	-0.5	0.67	0.015	2.	-0.5	0.67	0.015
1.	0	0.67	0.015	1.5	0	0.67	0.015	2.	0	0.67	0.015
1.	0.5	0.56	0.067	1.5	0.5	0.56	0.067	2.	0.5	0.46	0.20
1.	1.	0.46	0.20	1.5	1.	0.46	0.20	2.	1.	0.34	0.53
1.	1.5	0.36	0.49	1.5	1.5	0.36	0.49	2.	1.5	0.24	0.90
1.	2.	0.26	0.87	1.5	2.	0.26	0.87	2.	2.	0.23	0.93
1.	2.5	0.16	1.0	1.5	2.5	0.23	0.93	2.	2.5	0.14	1.0
1.	3.	0.16	1.0	1.5	3.	0.22	0.95	2.	3.	0.13	1.0
1.	3.5	0.18	0.99	1.5	3.5	0.22	0.95	2.	3.5	0.13	1.0
1.	4.	0.19	0.99	1.5	4.	0.22	0.95	2.	4.	0.19	0.99
1.	4.5	0.19	0.99	1.5	4.5	0.22	0.95	2.	4.5	0.19	0.99
2.5	-0.5	0.67	0.015	3.	-0.5	0.67	0.015	3.5	-0.5	0.67	0.015
2.5	0	0.57	0.058	3.	0	0.67	0.015	3.5	0	0.67	0.015
2.5	0.5	0.46	0.20	3.	0.5	0.46	0.20	3.5	0.5	0.56	0.058
2.5	1.	0.44	0.23	3.	1.	0.36	0.49	3.5	1.	0.44	0.23
2.5	1.5	0.34	0.53	3.	1.5	0.33	0.57	3.5	1.5	0.34	0.53
2.5	2.	0.24	0.90	3.	2.	0.33	0.57	3.5	2.	0.33	0.57
2.5	2.5	0.23	0.93	3.	2.5	0.33	0.57	3.5	2.5	0.33	0.57
2.5	3.	0.22	0.95	3.	3.	0.33	0.57	3.5	3.	0.33	0.57
2.5	3.5	0.22	0.95	3.	3.5	0.33	0.57	3.5	3.5	0.33	0.57
2.5	4.	0.22	0.95	3.	4.	0.33	0.57	3.5	4.	0.33	0.57
2.5	4.5	0.22	0.95	3.	4.5	0.33	0.57	3.5	4.5	0.33	0.57
4.	-0.5	0.67	0.015	4.5	-0.5	0.67	0.015				
4.	0	0.67	0.015	4.5	0	0.67	0.015				
4.	0.5	0.56	0.067	4.5	0.5	0.56	0.067				
4.	1.	0.46	0.20	4.5	1.	0.46	0.20				
4.	1.5	0.46	0.20	4.5	1.5	0.44	0.23				
4.	2.	0.36	0.49	4.5	2.	0.44	0.23				
4.	2.5	0.26	0.87	4.5	2.5	0.44	0.23				
4.	3.	0.26	0.87	4.5	3.	0.44	0.23				
4.	3.5	0.22	0.95	4.5	3.5	0.44	0.23				
4.	4.	0.22	0.95	4.5	4.	0.44	0.23				
4.	4.5	0.22	0.95	4.5	4.5	0.44	0.23				

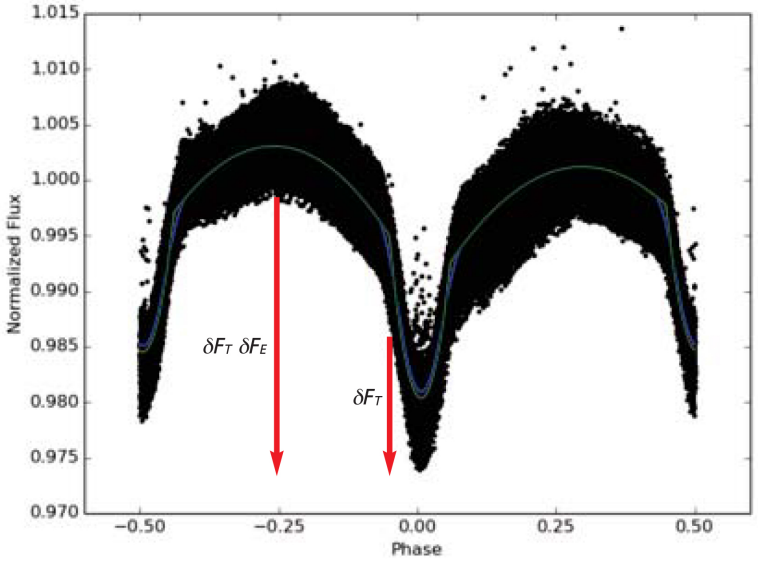


Fig. 7.14 Annotated light curve, KIC 6353203, morphology parameter = 0.64. δF_T is the change in flux due to the transit alone, $\delta F_T + \delta F_E$ is the change in flux due to both the transit and the ellipsoidal variations.

single long cadence integration (30 minutes). The ability to recognise that the signal comes from an eclipsing binary, and to measure the risk that it will be mis-identified as coming from a planet, is more important.

One way to measure this is through the morphology parameter used by the KEBC. As the morphology parameter increases, ellipsoidal variations become more evident and eclipse duration as a fraction of orbital period increases. However, there is no formal link between these two observable characteristics and the morphology parameter (Andrej Prsa, priv. comm.).

Coughlin et al. (2016) stated that only systems with a morphology parameter < 0.6 were included in *Kepler* DR24 and DR25. My preferred approach would have been to use δT (transit duration over orbital period), δF_T (the proportion of δF which is due to the transit) and $\delta F_T + \delta F_E$ (total δF, including both the eclipse and ellipsoidal variations), to derive the morphology parameter in each synthetic system (see Fig. 7.14). This would have matched the synthetic binary population to the observed population as recorded in NExSci more closely. Instead, a more empirical approach was required, which should always be regarded as subject to review as further data becomes available.

For the period distribution weighting, the KEBC, rather than NExSci, was utilised for the following reasons:

1. The KEBC includes binaries not included in NExSci, which lists only those binaries assigned a KOI number.

2. Systems with a stellar eclipse in NExSci may be blended or unblended. The KEBC has undertaken work to distinguish between blended and unblended binaries. In particular, in NExSci blended systems within the exclusion radius of the target cannot be distinguished as having an offset without follow-up work, such as that undertaken by the *Kepler* False Positive Working Group.

3. From DR25, grazing planets may carry the stellar eclipse flag if the resulting light curve is V-shaped.

4. Where there is a disagreement between NExSci and KEBC on orbital period in a given system, examination of the light curve indicates that more often than not the analysis from KEBC is correct. This occurs particularly in circular orbits with an observable secondary eclipse at $\phi = 0.5$: preliminary analysis in NExSci can indicate that these are differing depths of a single eclipse at half the true period. Such systems can be identified in NExSci catalogues through the use of the Odd/Even flag.

Since an empirical approach is required, I use cross validation in this experiment, dividing my simulated data into a training set and a testing set. The training set and the testing set are analysed separately. If the model produced with the training set fits the testing set, it is likely that the assumptions on which the analysis was based are correct. It was not possible to use this approach in deriving the intrinsic exoplanet distribution (Chapter 6), due to the small numbers of planets in many of the bins. However, since eclipsing binary simulations can be run relatively quickly utilising the synthetic *Kepler* field from Farmer et al. (2013), this approach was feasible in this part of the work.

A total of 40 unblended eclipsing binary synthetic populations were obtained, based on the *Kepler* field from Farmer et al. (2013). These synthetic populations were separated into two equal sets. Simulations 1-20 were allocated to Set 1 (the training set), and the simulations 21-40 were allocated to Set 2 (the testing set). Because it was always the intention to examine the impact of the weighting with υ (Section 7.2.2) at different periods, the numbers of binaries in which at least one component was a compact remnant were separated from the numbers of binaries in which neither component was a compact remnant. When weighting by υ was applied, binaries with at least one compact remnant ($\approx 0.4\%$) and binaries where $q > 1$ ($\approx 3\%$) were given a weighting of 1. The mean of the unblended binaries in each set, both before and after weighting by υ, was obtained.

Table 7.4 and Fig. 7.15 show the initial results of these simulations, comparing all observed binaries in the KEBC with those observed binaries with a morphology parameter < 0.7, separating out synthetic binaries with a compact component from those synthetic binaries without a compact component, and demonstrating the effect of weighting by υ, as described

Table 7.4 Orbital period distribution of binaries, KEBC vs simulations, no limit on eclipse duration, both eclipses detectable, mean of 20 simulations per set. Column 1: *log P*/days. Column 2: all KEBC entries. Column 3: KEBC entries with a morphology parameter < 0.7 (probably detached and probably semi-detached only). Column 4: mean number of binaries with a compact component, Simulations 1-20. Column 5: as column 4, Simulations 21-40. Column 6: mean number of binaries in which neither component is a compact remnant, before weighting by υ, Simulations 1-20. Column 7: As Column 6, Simulations 21-40. Column 8: as column 6, weighted by υ, based on $s = 2.5$, Simulations 1-20. Column 9: as column 8, Simulations 21-40.

log Period/ days	KEBC all	KEBC morph < 0.7	mean Set 1 Compact	mean Set 2 Compact	mean Set 1 Other All	mean Set 2 Other All	mean Set 1 Other Weighted by υ	mean Set 2 Other Weighted by υ
-0.4 to -0.2	251	72	0.35	1.05	285	289	256	261
-0.2 to 0.0	247	151	0.30	0.30	656	656	579	585
0.0 to 0.2	219	151	0.30	0.55	635	627	577	564
0.2 to 0.4	224	174	0.00	0.05	520	520	466	470
0.4 to 0.6	221	189	0.10	0.20	398	398	367	372
0.6 to 0.8	250	239	0.15	0.05	304	306	290	283
0.8 to 1.0	199	193	0.00	0.00	223	228	193	197
1.0 to 1.2	191	190	0.10	0.30	180	180	169	169
1.2 to 1.4	178	178	0.00	0.00	127	128	124	119
1.4 to 1.6	123	123	0.00	0.00	96	100	91	95
1.6 to 1.8	90	90	0.00	0.00	76	78	70	71
1.8 to 2.0	66	66	0.00	0.00	64	65	59	57
2.0 to 2.2	43	43	0.00	0.00	51	52	49	46
2.2 to 2.4	53	53	0.00	0.00	39	39	39	37
2.4 to 2.6	31	31	0.00	0.00	29	29	26	28
2.6 to 2.8	24	24	0.00	0.00	23	22	22	21
2.8 to 3.0	12	12	0.00	0.00	19	16	17	15
3.0 to 3.2	4	4	0.00	0.00	14	16	14	16

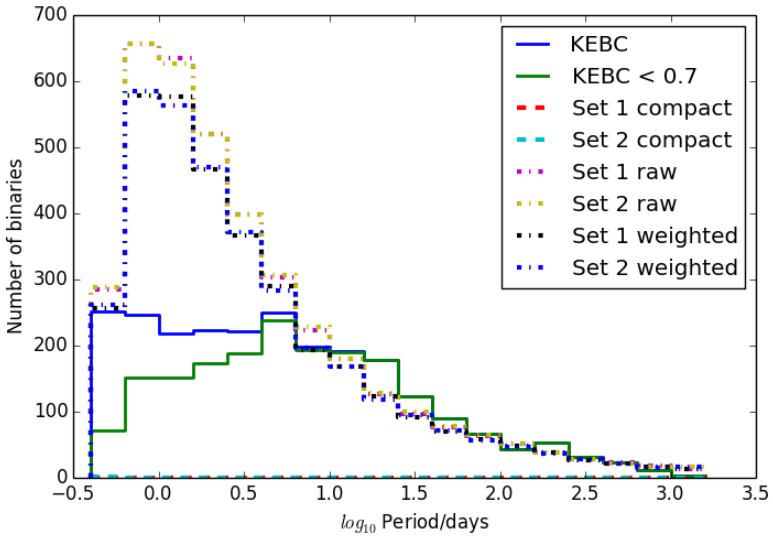


Fig. 7.15 Figure to accompany Table 7.4. No limit on the duration of the primary eclipse in grazing eclipsing binaries. KEBC: all KEBC entries. KEBC < 0.7: KEBC entries with a morphology parameter < 0.7. Set 1 compact: mean number of binaries with a compact component, Simulations 1-20. Set 2 compact: as Set 1 compact, Simulations 21-40. Set 1 raw: mean number of binaries in which neither component is a compact remnant, before weighting by υ, Simulations 1-20. Set 2 raw: As Set 1 raw, Simulations 21-40. Set 1 weighted: as column 6, weighted by υ, based on $s = 2.5$, Simulations 1-20. Set 2 weighted: as Set 1 weighted, Simulations 21-40.

Table 7.5 As Table 7.4, eclipse duration < 9% of orbital period in grazing binaries.

log Period /days	KEBC all	KEBC morph < 0.7	mean Set 1 Compact	mean Set 2 Compact	mean Set 1 Other All	mean Set 2 Other All	mean Set 1 Other Weighted by υ	mean Set 2 Other Weighted by υ
-0.4 to -0.2	251	72	0.35	1.05	134	133	89	86
-0.2 to 0.0	247	151	0.30	0.30	370	365	252	253
0.0 to 0.2	219	151	0.30	0.55	438	431	335	326
0.2 to 0.4	224	174	0.00	0.05	438	437	358	360
0.4 to 0.6	221	189	0.10	0.20	368	367	326	329
0.6 to 0.8	250	239	0.15	0.05	292	295	275	269
0.8 to 1.0	199	193	0.00	0.00	219	224	190	193
1.0 to 1.2	191	190	0.10	0.30	180	180	169	168
1.2 to 1.4	178	178	0.00	0.00	127	128	123	119
1.4 to 1.6	123	123	0.00	0.00	96	100	91	91
1.6 to 1.8	90	90	0.00	0.00	76	78	70	71
1.8 to 2.0	66	66	0.00	0.00	64	65	59	57
2.0 to 2.2	43	43	0.00	0.00	51	52	49	46
2.2 to 2.4	53	53	0.00	0.00	39	39	39	37
2.4 to 2.6	31	31	0.00	0.00	29	28	25	27
2.6 to 2.8	24	24	0.00	0.00	23	22	22	21
2.8 to 3.0	12	12	0.00	0.00	19	16	17	15
3.0 to 3.2	4	4	0.00	0.00	14	16	14	16

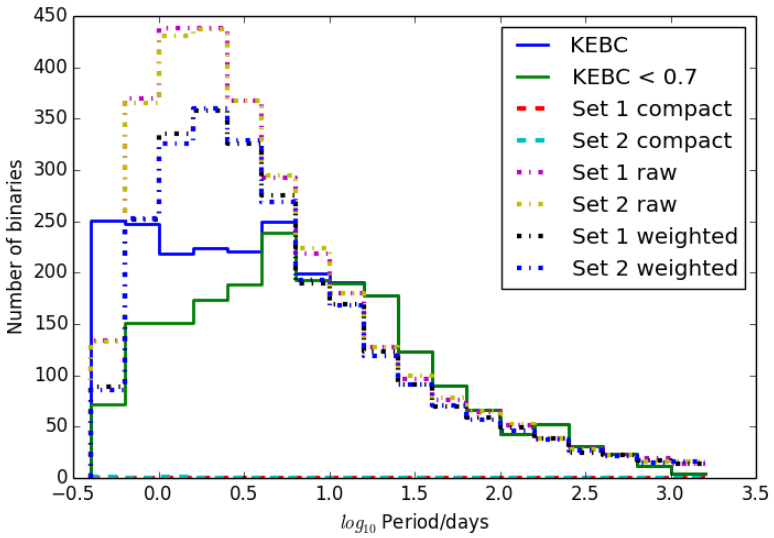


Fig. 7.16 Figure to accompany Table 7.5. As Fig. 7.15, with eclipse duration < 9% of orbital period in grazing binaries. The number of synthetic binaries with $log\ P < 1.0$ days is reduced.

in the previous section, with $s = 2.5$. In Table 7.4 and Fig. 7.15 no maximum limit was set on δT.

Table 7.5 and Fig. 7.16 demonstrate the effect of limiting δT in grazing eclipsing binaries to 9%. No limit was applied to non-grazing binaries (ie those with $b < R_1 - R_2$) where R_1 and R_2 are the radii of the two stars. At short periods, fewer binaries are included in Table 7.5 and Fig. 7.16 than in Table 7.4 and Fig. 7.15, and weighting by υ further reduces the numbers of synthetic binaries counted, but in Table 7.5 and Fig. 7.16 there is still an excess of binaries where synthetic binaries are weighted by υ with $s = 2.5$ at $log\ P \lessapprox 0.8$ days, an approximate match at $0.8 \lessapprox log\ P/\text{days} \lessapprox 1$, a deficit at $1 \lessapprox log\ P/\text{days} \lessapprox 1.8$, and an approximate match at $log\ P \gtrapprox 1.8$ days (unweighted). Weighting by υ appears to make little difference at $log\ P \gtrapprox 1.2$ days.

The training set and testing set appear to be in agreement with each other.

Reasons for the excess at $log\ P \lessapprox 0.8$ days may include the following:

1. Simulated eclipses may be identified which, due to a shallow eclipse depth and large ellipsoidal variations, are not identifiable in the observed population (ie issues arising from completeness). Note that non-eclipsing ellipsoidal binaries may be present in the observed population, but are excluded from the synthetic population.

2. Systems cannot normally form at $a < 1$ AU (Moe and Di Stefano, 2017) and must migrate to that configuration, but a flat initial period distribution, as utilised by BiSEPS, produces an excessive number of short period systems, reflected in an excess in the number of short period eclipsing binaries. Fig. 7.17 demonstrates this in a 1 deg^2 field centred on $l = 50°$, $b = 8°$, the same field as was used to weight υ at $s = 0$. Panel (a) illustrates all solar metallicity binaries in the field, and panel (b) illustrates the number of solar metallicity eclipsing binaries in a sample simulation. There are significant differences in the period distribution in the two panels.

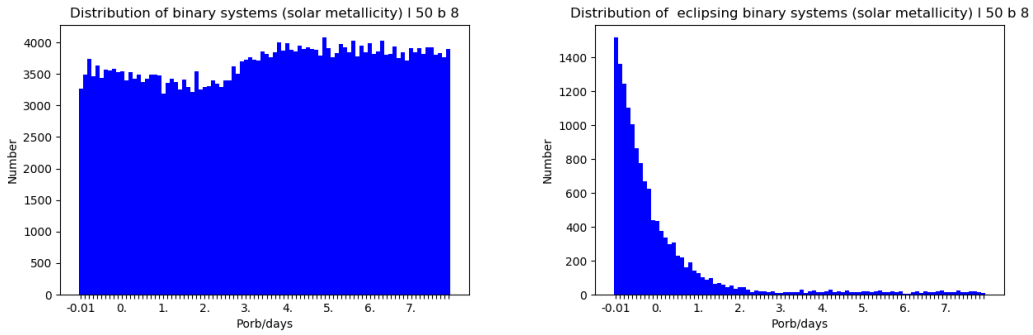
Reasons for the small deficit at $1 \lessapprox log\ P/\text{days} \lessapprox 1.8$ may include the following:

1. Eccentricity: this is currently not incorporated in BiSEPS and is a subject for future work. At $P > 10$ days, the assumption of low eccentricity is no longer generally valid.

2. Mis-characterisation in the observed population of planets, particularly giant planets, or brown dwarfs, as binaries, possibly due to the shape of the light curve.

3. A "pile up" of binaries, that is an excess in the real binary distribution due to a hypothetical slowing down in period decrease.

To explore this deficit further, the condition in Table 7.4 and Table 7.5, Fig. 7.15 and Fig. 7.16, that both eclipses should be detectable, was relaxed. The distribution this change produced are presented in Table 7.6 and Fig. 7.18. This has the obvious effect of increasing the number of binaries that can be detected at all orbital periods, and reflects the conditions



(a) All solar metallicity binaries (b) Solar metallicity eclipsing binaries

Fig. 7.17 Comparing the number of of solar metallicity binaries with eclipsing binaries in a 1 deg^2 synthetic field. This is the same field that was used to weight υ at $s = 0$. Note the different y axes.

under which data would be extracted from the full PLATO simulations. It reduces the deficit referred to above, but at the same time increases the excess at the shortest periods.

Using the data from Table 7.6 and Fig. 7.18, the next step was to find a correction factor for each period bin which would allow the synthetic population to approximate the observed population with a morphology parameter < 0.7, without overfitting the data. The correction factor represents a smoothed curve, binned logarithmically by period. This was found to reproduce a smoothed KEBC distribution to within 15%, apart from the bin $-0.2 < log\ P/\text{days} < 1$, which appears to have an anomalously large number of binaries in the observed population.

Table 7.7 and Fig. 7.19 illustrate the results of the calibration exercise, comparing a fully calibrated synthetic population with the observed population from the KEBC with a morphology parameter < 0.7. As can be seen, the fully calibrated synthetic population is not an exact match for the observed population, but it is much closer than in previous figures and tables. The synthetic population when calibrated in this way rises to a peak at $0.6 < log\ P/\text{days} < 0.8$ and then falls again, relatively smoothly. Weighting by υ with $s = 2.5$ was required up to and including the bin $1.2 < log\ P/\text{days} < 1.4$: at longer periods, weighting by υ was not required.

Fig. 7.20 illustrates the period calibration factor, also given in Table 7.7, Column 4. This rises to a maximum of 1.25 at $1.2 < log\ P/\text{days} < 1.4$ before setting at 1 for $log\ P > 1.6$ days. There is an additional peak in the bin $-0.4 < log\ P/\text{days} < -0.2$ because this is a synthetic population, so no binaries with sinusoidal light curves that are not actually transiting would be included.

Table 7.6 As Table 7.5, but the condition that both eclipses must be detectable has been relaxed. This increases the number of observable binaries at all orbital periods.

log Period days	KEBC all	KEBC morph < 0.7	mean Set 1 Compact	mean Set 2 Compact	mean Set 1 Other All	mean Set 2 Other All	mean Set 1 Other Weighted by υ	mean Set 2 Other Weighted by υ
-0.4 to -0.2	251	72	2.85	3.55	164	166	93	91
-0.2 to 0.0	247	151	1.45	1.80	457	451	262	262
0.0 to 0.2	219	151	1.70	1.90	508	501	344	334
0.2 to 0.4	224	174	1.15	1.40	493	497	364	367
0.4 to 0.6	221	189	0.55	0.85	417	415	330	334
0.6 to 0.8	250	239	0.65	0.40	324	327	278	273
0.8 to 1.0	199	193	0.50	0.80	247	252	193	196
1.0 to 1.2	191	190	0.30	0.35	203	205	171	171
1.2 to 1.4	178	178	0.40	0.25	143	146	126	121
1.4 to 1.6	123	123	0.15	0.00	110	113	93	95
1.6 to 1.8	90	90	0.00	0.00	87	90	71	73
1.8 to 2.0	66	66	0.00	0.00	73	75	60	60
2.0 to 2.2	43	43	0.00	0.00	59	59	49	48
2.2 to 2.4	53	53	0.00	0.00	45	46	41	40
2.4 to 2.6	31	31	0.00	0.00	32	33	28	30
2.6 to 2.8	24	24	0.00	0.00	28	27	24	23
2.8 to 3.0	12	12	0.00	0.00	23	20	19	17
3.0 to 3.2	4	4	0.00	0.00	18	19	15	17

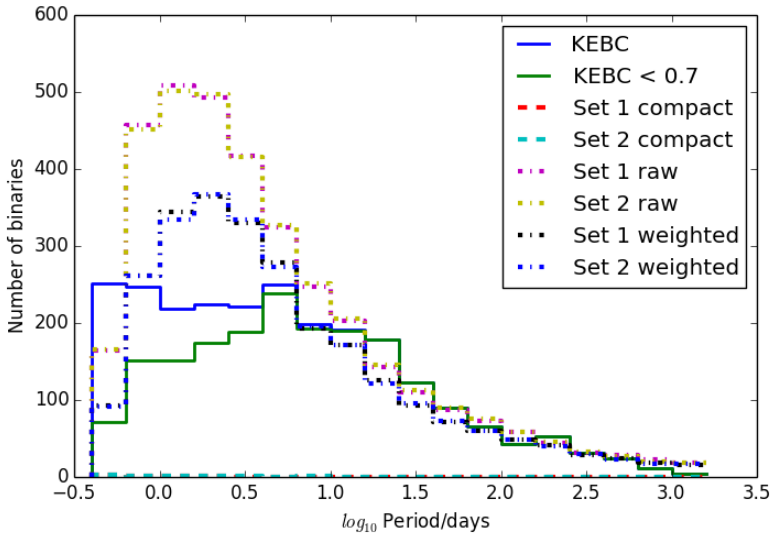


Fig. 7.18 Figure to accompany Table 7.6. As Fig. 7.16, but the condition that both eclipses must be detectable has been relaxed. This increases the number of observable binaries at all orbital periods.

Table 7.7 Final period calibration. Column 1: log Period/days. Column 2: Eclipsing binaries in the KEBC with a morphology parameter < 0.7 (probably detached and probably semi-detached only). Column 3: indicates if calibration by υ with $s = 2.5$ was used in this period bin. Column 4: The factor by which the mean of the number of binaries in a set was multiplied in this period bin. Column 5: mean of the number of binaries in which neither component is a compact remnant, weighted by υ based on $s = 2.5$ and corrected by the period calibration coefficient, Simulations 1-20. Column 6: as column 5, Simulations 21-40.

log Period/ days	KEBC morph < 0.7	IMR weighted Y/N	Period correction factor	mean Set 1	mean Set 2
-0.4 to -0.2	72	Y	0.75	70	68
-0.2 to 0.0	151	Y	0.40	105	105
0.0 to 0.2	151	Y	0.45	155	150
0.2 to 0.4	174	Y	0.50	182	183
0.4 to 0.6	189	Y	0.67	221	223
0.6 to 0.8	239	Y	0.90	251	246
0.8 to 1.0	193	Y	1.00	193	196
1.0 to 1.2	190	Y	1.11	189	190
1.2 to 1.4	178	Y	1.25	157	151
1.4 to 1.6	123	N	1.11	122	125
1.6 to 1.8	90	N	1.00	87	90
1.8 to 2.0	66	N	1.00	73	75
2.0 to 2.2	43	N	1.00	59	59
2.2 to 2.4	53	N	1.00	45	46
2.4 to 2.6	31	N	1.00	32	33
2.6 to 2.8	24	N	1.00	28	27
2.8 to 3.0	12	N	1.00	23	20
3.0 to 3.2	4	N	1.00	18	19

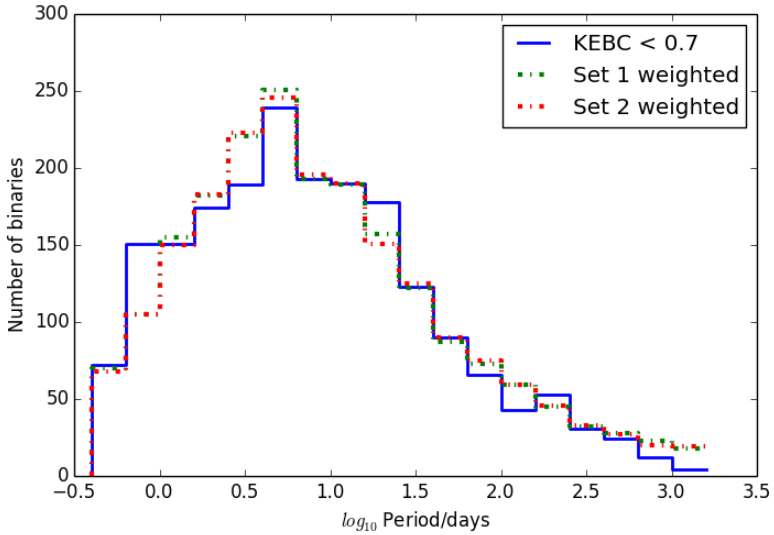


Fig. 7.19 Figure to accompany Table 7.7. KEBC < 0.7: KEBC entries with a morphology parameter < 0.7. Set 1 weighted: mean of binaries in which neither component is a compact remnant, weighted by υ, based on $s = 2.5$ and corrected by period, Simulations 1-20. Set 2 weighted: as Set 1 weighted, Simulations 21-40.

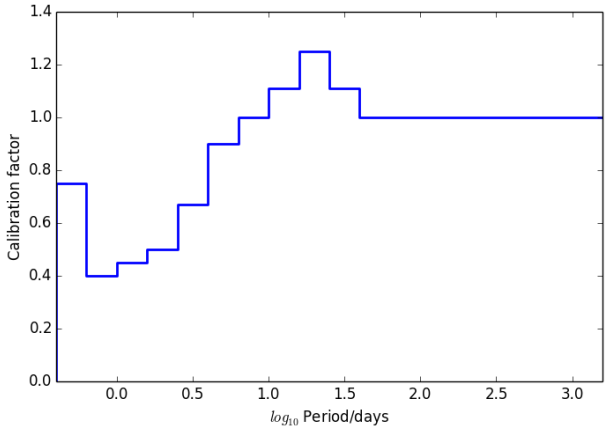


Fig. 7.20 Final period correction factor, by period bin. To accompany Table 7.7. This was found to reproduce a smoothed KEBC distribution to within 15%, apart from the bin $-0.2 <$ *log P*/days < 1, which has an anomalously large number of binaries in the observed population.

This calibration is approximate, and is expected to be improved in future work, in particular once Gaia DR3 is available, probably in 2019. This is the release that is expected to include the full binary catalogue. The question of whether planets are being mischaracterised as binaries at $\approx$ 10-40 days can be further addressed at that point. In addition, eccentricity and the possibility of a binary "10-40 day pile-up" are issues that should be addressed in future work.

7.4 Application to the PLATO fields

The synthetic binary sample presented here represents those binaries likely to be observed in a transiting exoplanet survey such as *Kepler* and PLATO, not the intrinsic eclipsing binary population of the Galaxy.

The samples are subject to significant uncertainties in the free parameters of the population model, which future work should be able to resolve more closely. These binary star estimates are intended to be used with the exoplanet estimates from Chapter 6 to predict the ratio of planets to blended and unblended eclipsing binaries (Chapter 8).

The samples presented here are the mean of five full simulations, interpreted with PLATO precision in a one hour integration. The mean total number of eclipsing binaries per simulation is 3.13 x 10^6 in LLS, and 8.53 x 10^6 in LLN: in other words, LLS has 37% of the total number of binaries simulated in LLN. The mean distribution as a function of Galactic longitude, l, is presented in Fig. 7.21. The distribution by l is governed by (i) relationship to the Galactic Centre, (ii) extinction, derived from Drimmel et al. (2003), and (iii) the number of binaries seeded in the model Galaxy as a function of l and b (Chapter 4). The region in LLN 81.5° $\leq l \leq$ 89.5° has a similar angular separation from the Galactic centre as the region in LLS 278.5° $\leq l \leq$ 270.5°. The mean number of eclipsing binaries simulated in these regions are 6.18 x 10^5 (LLN) and 8.46 x 10^5 (LLS). So, the population in LLS is 137% of that in LLN at similar angular separation from the Galactic centre. While simulating the wider field described in Chapter 4, I monitored the numbers of stars at equivalent l and at equivalent $|b|$ (above and below the Galactic plane), and found that the number of systems simulated in the Southern hemisphere was generally about 111% of the number of systems simulated in the Northern hemisphere. Therefore, this excess in LLS at equivalent angular separation from the Galactic centre is as expected.

I analysed the synthetic eclipsing binaries at three angular separations from synthetic potential PLATO target stars, interpreting the simulations with PLATO precision in a one hour integration (Chapter 5).

(a) Unblended binaries (UB, 0.0"). These binaries are included under the following assumptions: (i) The binary is the target observed by PLATO. (ii) It is possible to disentangle

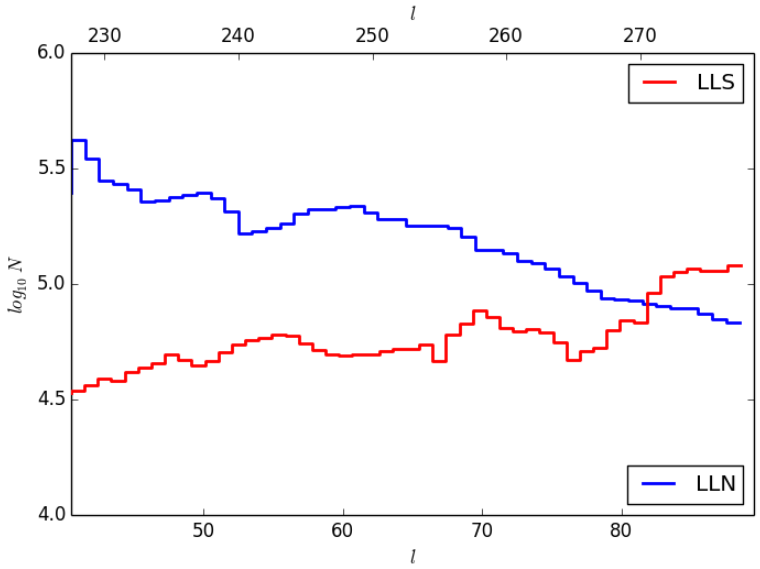


Fig. 7.21 Mean number of eclipsing binaries per simulation, as a function of Galactic longitude l, before interpretation with PLATO precision. The distribution by l is governed by (i) relationship to the Galactic Centre, (ii) extinction, derived from Drimmel et al. (2003), and (iii) the number of binaries seeded in the model Galaxy as a function of l and b (Chapter 4). The top x axis gives l for LLS, and the bottom$x$ axis gives l for LLN.

all other information and consider the true depths of the binary eclipses. It is recognised that assumption (ii) is unlikely to be fulfilled in actual PLATO observations in more densely populated regions of the Galaxy, although where the target dominates the flux in its pixel this disentangling should be easier to obtain.

(b) Binaries within the same pixel (7.5"). Blending within the same pixel will in practice be a lower limit on the distance at which blending will occur in PLATO observations. These binaries are identified on the assumption that the target is positioned at the centre of the pixel.

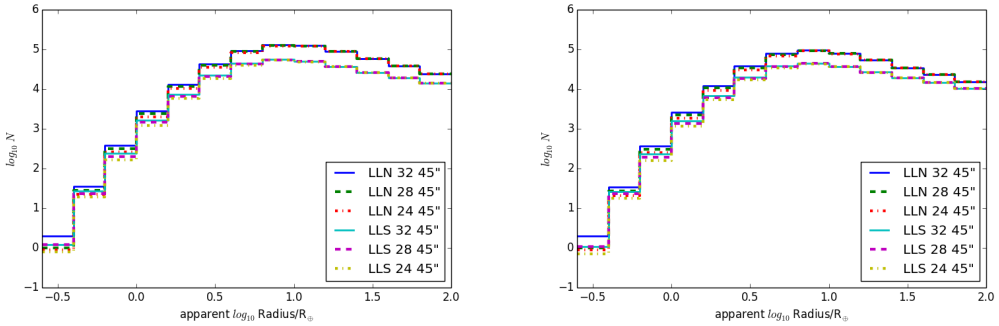
(c) Binaries within the same imagette (45.0"). Blending within the same imagette will in practice be the upper limit on the distance at which blending with the full PSF of the blended eclipsing binary will occur in PLATO observations, assuming a 6 pixel x 6 pixel imagette with the target at the centre.

In practice, blending in PLATO observations will fall between these two limits, 7.5" and 45.0". The samples presented here are intended to provide approximate minimum and maximum levels of contamination, not to be firm predictions in their own right. Comparison with the unblended sample is also beneficial in understanding the effects of contamination in binaries mimicking exoplanets of a given planet class, although it is not expected that, in practice, the unblended sample will be clearly observed.

7.4.1 Within the same imagette, < 45"

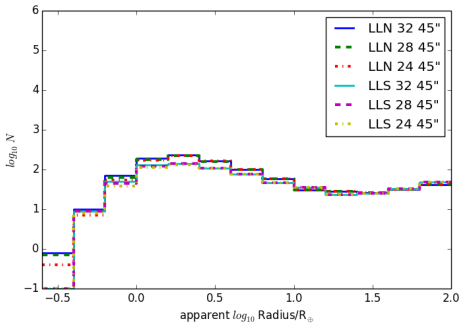
The binary sample within 45" of the target is intended to include all eclipsing binaries whose PSF may be blended with the PSF of any of the target stars. The wings of the PSFs of more distant binaries may also be blended with the PSF of the target, in a similar fashion to the false positives flagged as ephemeris matches in the *Kepler* data. As noted in Section 7.1, Coughlin et al. (2014) recorded ephemeris matches in *Kepler* data where the parent is up to 1,000" away from the child, with most parents between 5"-400" from the child: this is in the context of a *Kepler* plate scale of 3.98"/pixel. PLATO's plate scale is 15.0"/pixel. The circumstances which lead to an ephemeris match often include alignment along rows or, more frequently, columns. A small misalignment of the cameras may reduce the risk of ephemeris matches due to charge transfer in more than one camera: however, as described in the PLATO Definition Study Report, misalignment may also reduce the overlap of the cameras and reduce the number of stars that can be observed by all cameras in a group. The target given in the Report is that misalignment should not exceed ≈ 3.4'.

A comprehensive catalogue for use with PLATO, listing all binaries and variable stars within the field, not just those which may be observed, is therefore essential. It is assumed that this tool will be available from forthcoming Gaia data releases, and will eliminate blends with just the wings of the PSF of a nearby star from further consideration.

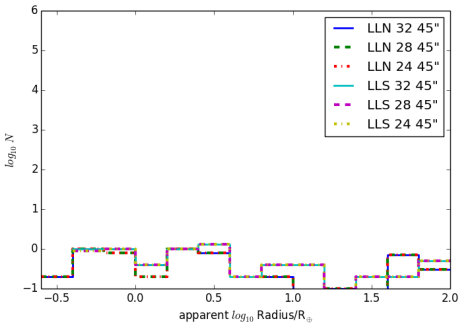


(a) $V \leq 26$

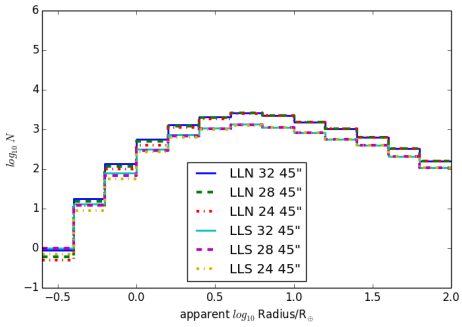
(b) $V \leq 16$, F5-K



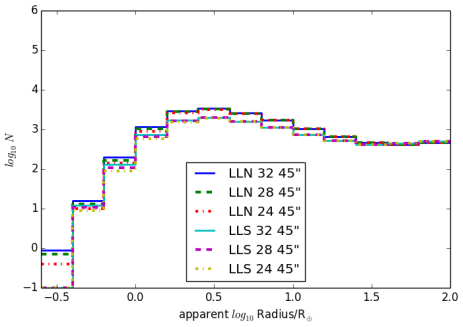
(c) P1



(d) P2



(e) P4



(f) P5

Fig. 7.22 Synthetic binaries within 45" of the target, as a function of apparent radius. LLN is Long Look North, LLS is Long Look South, and the number in the legend refers to the number of cameras. Binaries with an apparent radius $-0.4 < log\ R/R_\oplus < 0.2$ mimic terrestrial planets, binaries with an apparent radius $0.2 < log\ R/R_\oplus$ 0.8 mimic Neptune-like planets, binaries with an apparent radius $0.8 < log\ R/R_\oplus < 1.4$ mimic giant planets, and binaries with an apparent radius $log\ R/R_\oplus > 1.4$ are clearly stellar in nature. In P2, the blended population sample is dominated by small number statistics.

Table 7.8 Apparent radius of binaries mimicking exoplanets by type of exoplanet. See Section 6.3 for more information on the planet classes.

log Radius/$R_\oplus$	Planet type
-0.6 to -0.4	Smaller than considered in the exoplanet population
-0.4 to -0.2	Mars-analogue
-0.2 to 0.0	Earth-analogue (1)
0.0 to 0.2	Earth-analogue (2)
0.2 to 0.4	Small Neptunes (1)
0.4 to 0.6	Small Neptunes (2)
0.6 to 0.8	Large Neptunes
0.8 to 1.0	Giant Planets (1)
1.0 to 1.2	Giant Planets (2)
1.2 to 1.4	Giant Planets (3)
1.4 to 1.6	Stellar
1.6 to 1.8	Stellar
1.8 to 2.0	Stellar

Fig. 7.22 presents the mean of five simulations and covers the separation up to 45". The figures show the distribution over apparent radius, ie the radius of a planet that the binary is mimicking. A detailed breakdown is included in Appendix D, Tables D5, D6, D11, D12, D17, D18, D23, D24, D29, D30, D35 and D36. These are reported in units of $R_\oplus$ to enable easy comparison with the exoplanet simulations. For easy reference, the Sun at 1 $R_\odot$ is $\approx$ 100 $R_\oplus$. Detailed information relating radius to planet type is included in Section 6.3, and is summarised in Table 7.8.

Fig. 7.22 shows a breakdown by field (LLN and LLS); by number of cameras, demonstrating the effect of reducing the number of cameras from 32 (8/group) to 24 (6/group); and by PLATO priority population, with two additional populations included: $V \leq 16$, F5-M, which approximates the population observed by *Kepler*, and $V \leq 26$, in other words everything PLATO detectors are sensitive to, whether or not the system is in a PLATO priority population. The two populations $V \leq 26$ and $V \leq 16$, F5-M, are included for better comparison with data from other surveys, such as *Kepler*.

Approximately twice as many binaries within the same imagette are detected in the LLN field than in LLS, which is consistent with the ratio of systems in LLN to LLS (see Section 6.4.1) and with the total number of eclipsing binaries simulated before interpretation with PLATO precision (Fig 7.21).

In P2 (Fig. 7.22, panel (d)), small number statistics make trends impossible to distinguish. At $V \leq 8$, these target stars are essentially rare.

In P1, there is a marked peak at $0.0 \lesssim log\ R/R_\oplus \lesssim 0.6$ in both Long Look fields (Fig. 7.22, panel (c)), mimicking terrestrial and Neptune-like planets. The population falls to a minimum at $1.2 \lesssim log\ R/R_\oplus \lesssim 1.6$, before rising slowly as the apparent radius approaches Sun-like values ($log\ R/R_\oplus$ 2.0). Terrestrial and Neptune-like planets are expected to be of most interest to PLATO and P1 is the population most of interest, as these are the stars for which asteroseismology will be routinely completed, better constraining stellar properties and therefore the radius and age of the planet. It is therefore useful to be aware that at maximum blending, the planets most likely to be of interest are also likely to be the most contaminated by background eclipsing binaries.

In the more general populations, which include systems that PLATO is capable of detecting but is not prioritising, the maximum is at $0.6 \lesssim log\ R/R_\oplus \lesssim 1.2$ (Fig. 7.22, panel (a) ($V \leq 26$), Fig. 7.22, panel (b) (*Kepler*-like: $V \leq 16$, F5-M)). The binaries near this maximum would mimic large Neptunes and giant planets.

In P5 (Fig. 7.22, panel (f)), the maximum lies at $0.2 \lesssim log\ R/R_\oplus \lesssim 0.8$, mimicking Neptune-like planets. The limiting magnitude is fainter than P1 and brighter than P4 (Fig. 7.22, panel (e)), where the peak in the distribution lies at $0.4 \lesssim log\ R/R_\oplus \lesssim 1.0$, mimicking Neptune-like planets and smaller giant planets.

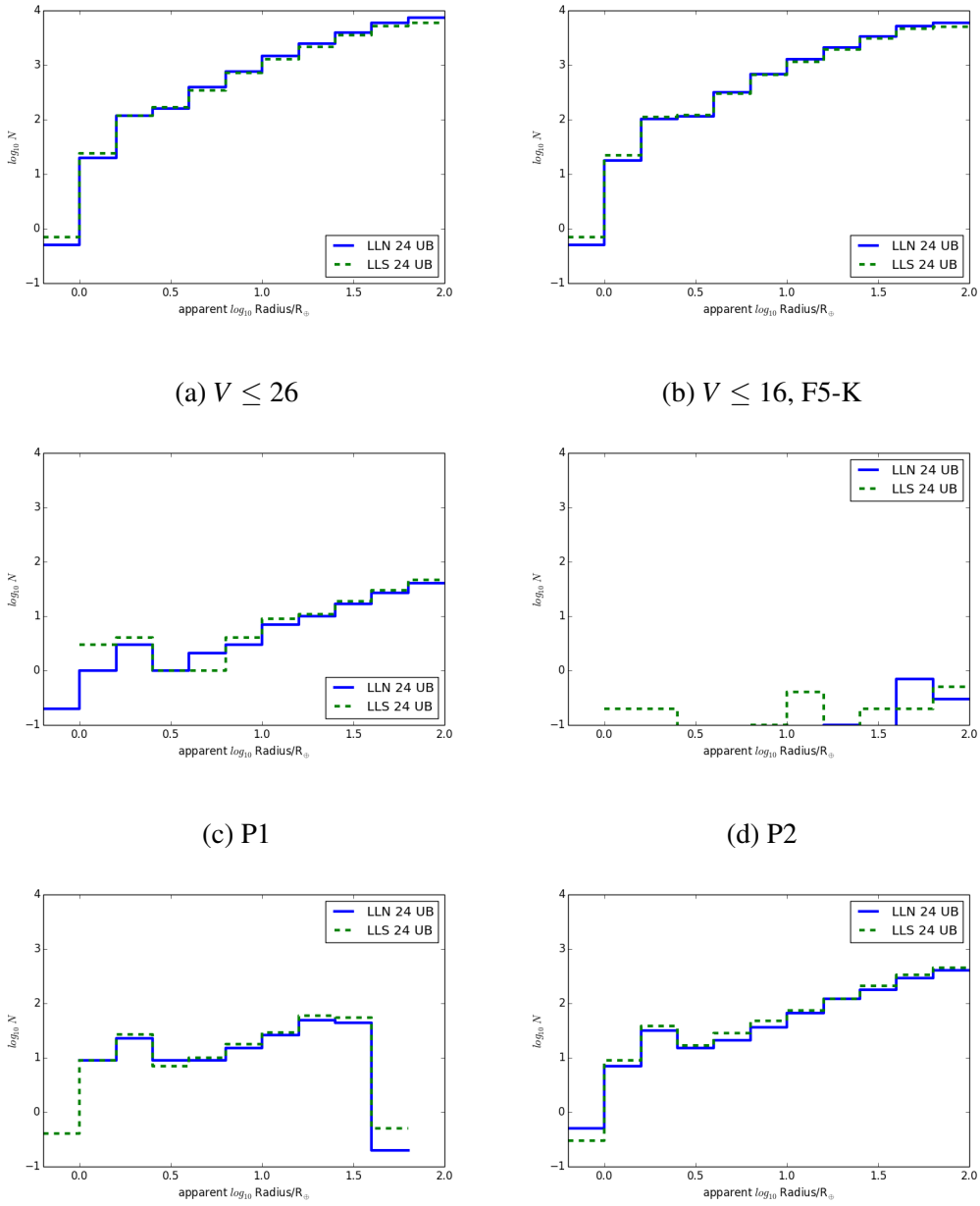
When blends are considered out to 45", the limiting magnitude in the PLATO priority classes is a key factor in determining which type of planets are most likely to be mimicked by background eclipsing binaries.

7.4.2 Unblended

Results for unblended binaries are presented in Fig. 7.23 and Appendix D, Tables D1, D2, D7, D8, D13, D14, D19, D20, D25, D26, D31 and D32.

In the unblended populations, there is a clear distinction between the population dominated by binaries in which one component is a WD, mimicking planets with an apparent radius of $0.0 \lesssim log\ R/R_\oplus \lesssim 0.4$ (Table 7.9), and the grazing population, which may mimic planets of any size. Given that one of the constraints on the BiSEPS simulation is a stellar radius of $\geq$ 0.1 $R_\odot$ at ZAMS, any unblended eclipsing binary with no compact component mimicking a planet of $\lesssim$ 10 $R_\oplus$ is grazing.

The apparent radius is derived from the fractional decrease in flux δF during the eclipse. Section 1.5, Equations 1.10, 1.11 and 1.12 summarise how δF is derived in the case of an eclipsing binary, where both objects are self luminous, although limb darkening (Section 3.1.1) must also be taken into account; and Eq. 1.1 demonstrates how δF is used, in conjunction with an estimate of the stellar radius, to derive an estimate of the planetary radius.



(a) $V \leq 26$

(b) $V \leq 16$, F5-K

(c) P1

(d) P2

(e) P4

(f) P5

Fig. 7.23 Unblended synthetic binaries, as a function of apparent radius. LLN is Long Look North, LLS is Long Look South, and the number in the legend refers to the number of cameras. In P1, P4 and P5, the maximum at $0.2 \lesssim log\ R/\mathrm{R}_\oplus \lesssim 0.4$ is dominated by binaries in which one object is a white dwarf (Table 7.9). In P2, the sample is dominated by small number statistics. In P4, the sharp drop at $R \approx 1.6\ \mathrm{R}_\oplus$ is determined by the maximum radius of an M-class star.

Table 7.9 Proportion of unblended binaries mimicking planets of a given radius which include a WD component, by PLATO priority class and by field. No unblended binaries were noted mimicking planets $log\ R/R_\oplus < -0.2$, and there are no unblended binaries in which one component is a WD mimicking planets $log\ R/R_\oplus > 0.8$. n/a means no binaries in this PLATO priority class mimicking planets of this radius in this field.

$log\ R/R_\oplus$	P1 LLN	P1 LLS	P2 LLN	P2 LLS	P4 LLN	P4 LLS	P5 LLN	P5 LLS
-0.2 to 0.0	0.0	n/a	n/a	n/a	n/a	0.0	0.20	0.50
0.0 to 0.2	1.0	0.85	n/a	1.0	0.89	0.91	0.82	0.77
0.2 to 0.4	0.85	0.91	n/a	1.0	0.90	0.88	0.87	0.83
0.4 to 0.6	0.0	0.14	n/a	0.0	0.18	0.28	0.18	0.17
0.6 to 0.8	0.0	0.0	n/a	n/a	0.0	0.0	0.0048	0.00

In P1 (Fig. 7.23, panel (c)) and in P5 (Fig. 7.23, panel (f)), binaries in which one object is a WD dominate the synthetic population up to $log\ R/R_\oplus = 0.4$. Table 7.9 quantifies this, by PLATO priority population and by apparent radius. At $0.0 < log\ R/R_\oplus < 0.4$, binaries in which one component is a white dwarf make up $\gtrsim 80\%$ of the unblended binary population. Grazing binaries are becoming more significant at apparent radii above $log\ R/R_\oplus = 0.4$ and there are no compact objects in binaries mimicking planets above $log\ R/R_\oplus = 0.8$. Indeed, the lack of a contribution by compact binaries at larger apparent radii emphasises the point that where in blended populations blended compact binaries appear to mimic larger planets, this is a natural result of the fact that the apparent radius is being assessed not against the true radius of the other star in the binary (and ignoring the fact that both stars are self-luminous, which a planet is not), but against the radius of the target star.

In P4 (Fig. 7.23, panel (e)) there are again two maxima, one dominated by binaries in which one object is compact ($0.2 < log\ R/R_\oplus < 0.4$, Table 7.9) and one dominated by other binaries. According to our models, binaries with one M-class star and one CO WD and with periods short enough to be included in this analysis are the surviving remnants of a common envelope phase, and therefore are interesting objects for study in their own right. In a long period M-class MS/WD binary both components have evolved independently: however, since the period in this latter group is invariably longer than the upper limit for data extracted from the simulations and presented in this sample, $log\ P/days = 3.2$, these objects are not included and would, in any case, have a period too long to be considered for follow up in a PLATO Long Look field, as the requisite three detections would not be obtained. Systems which have survived a common envelope phase as a binary naturally have very short periods.

Note that, in P4, there are no unblended binaries with an apparent radius $log\ R/R_\oplus > 1.8$: this is a natural consequence of the maximum radius of an M-class star.

7.4.3 Comparison: blended and unblended binaries

In the PLATO priority classes, more synthetic unblended eclipsing binaries are detectable in LLS than LLN at all apparent radii. This is despite the fact that, as shown in Fig. 7.21, there are more eclipsing binaries simulated in LLN than in LLS. This is also in contrast to blended eclipsing binaries within the same imagette, where more synthetic binaries were detectable in LLN than in LLS in the P1, P4 and P5 populations (Section 7.4.1). In the two background samples, $V < 26$ and $V < 16$, F5-M, at smaller apparent radii more unblended eclipsing binaries are detected in LLS, but at larger radii more binaries are detected in LLN.

These results are interesting when compared with the results from Chapter 6 concerning planets, where more terrestrial planets were detected in LLS, and more giant planets were detected in LLN.

As discussed in Section 6.4.1, more planets in total were simulated in LLN than LLS but, when analysed by PLATO priority population, more planets were observable in LLS than LLN. It was speculated in Section 6.4.1 that this may be because planets in the PLATO priority populations are more likely to be nearby than in the more general populations, and to observe LLS it is necessary to look through the Galactic plane, where the stellar density is likely to be greatest. The total number of binaries in the LLN synthetic field exceeds the total number of binaries in the LLS synthetic field but, in the PLATO priority populations, more unblended binaries are observed in LLS than in LLN.

To investigate the relationship between distance and detection of a blended or unblended binary mimicking a planet of a given radius, distance v apparent radius plots were obtained for both the unblended and blended binary samples, in both LLN and LLS, for the P5 population, observed by 24 cameras (6 cameras/group), one hour integration. These are uncalibrated in the sense that there has been no attempt to calibrate by detection probability υ as described in Section 7.2, or by initial period distribution as described in Section 7.3. The unblended population is illustrated in Fig. 7.24, panels (a) and (b), and the blended population within the same pixel (7.5") in panels (e) and (f) (see Section 7.4.4). The distribution by apparent radius against distance is similar whether blending within the pixel or blending within the imagette is considered.

Fig. 7.24 shows that for the P5 population observed with 24 cameras in both LLN and LLS, unblended binaries mimicking giant planets are detected at similar distances to such planets, up to ≈ 1 kpc. Fig. 6.16 and Fig. 6.17 have been reproduced as panels (c) and (d) in Fig. 7.24 for easy reference. Binaries mimicking terrestrial planets can be detected at greater distance than true terrestrial planets. There are some outliers but, generally, unblended eclipsing binaries mimicking giant planets may be detected at greater distances than unblended binaries

mimicking terrestrial planets, just as giant planets can be detected at greater distances than terrestrial planets.

Fig. 7.24, panels (e) and (f), show that, for a sample where the targets are P5 stars observed with 24 cameras, blended binaries are detected at much more significant distances of up to 10 kpc in a one hour integration, with the furthest binaries being those mimicking terrestrial planets. This is as expected, as blended binaries mimicking terrestrial planets may also be the most heavily blended.

Fig. 7.24 suggests that two factors will become important when the ratio of planets to binaries is considered in Chapter 8: the stellar density of the background, and whether or not one is viewing through the Galactic plane. Extinction will also vary across the sky depending on local conditions (Fig. 7.21), but over such large fields is unlikely to be solely responsible for the observed trends. See Section 8.2.2 for further discussion of the stellar density along the line of sight to the LLN and LLS fields.

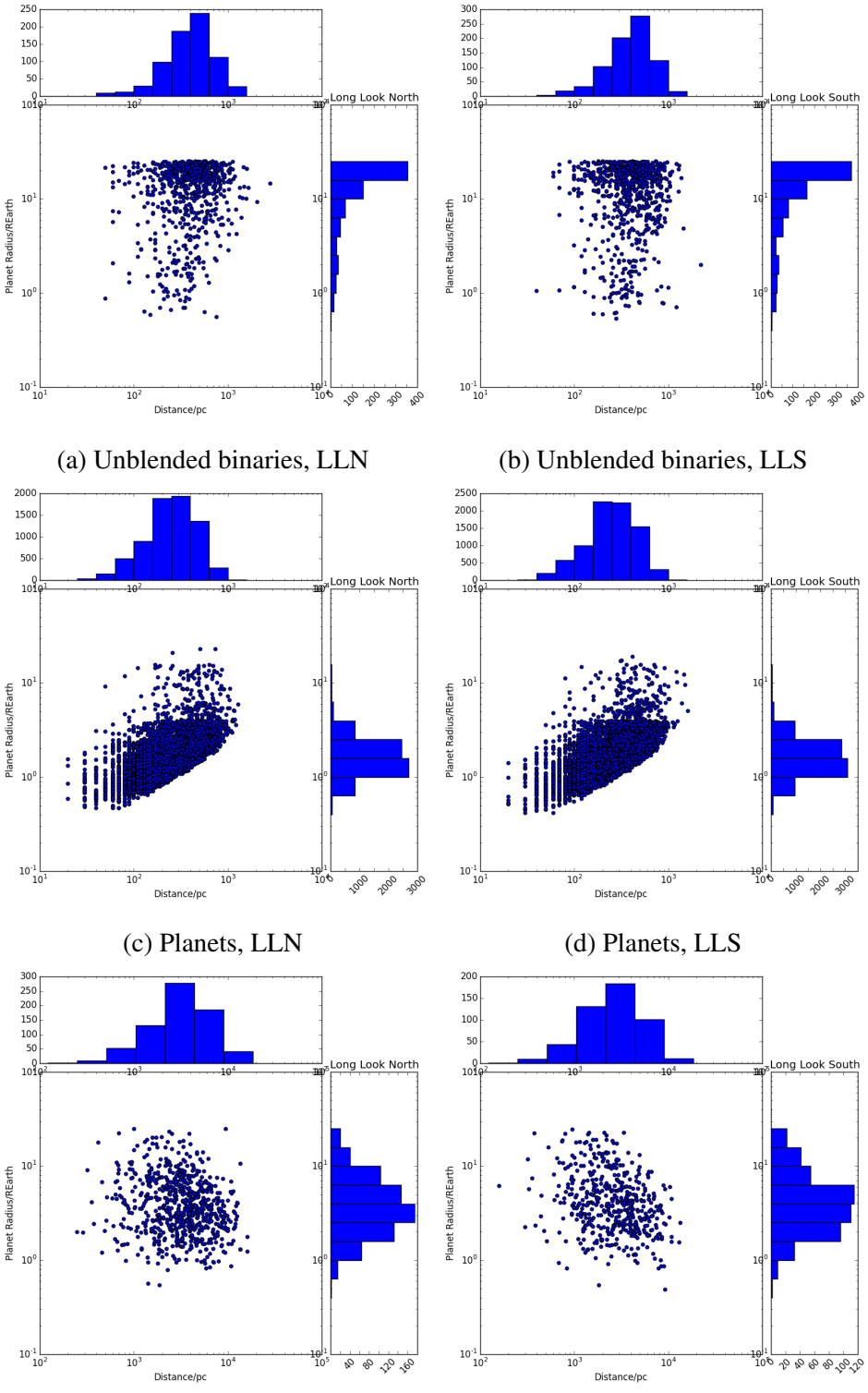
Fig. 7.25 compares, in the P1 population in the LLN field, the number of unblended eclipsing binaries with the number of blended eclipsing binaries within the same imagette, observable by 24 cameras. Although both samples show a maximum at $0.2 < log\ R/R_{\oplus} < 0.4$, in the unblended population this is dominated by binaries in which one object is compact (Section 7.4.2), but in the blended population the apparent radius is that recorded *after* dilution by blending: the true eclipse depth is deeper. Approaching solar radius, the unblended and blended samples converge.

7.4.4 Within the same pixel, < 7.5"

Fig. 7.26 and Fig. 7.27 compare blended and unblended binaries within 45" of the target (same imagette), blended and unblended binaries within 7.5" (same pixel) and cases where the target is an unblended binary. Contamination by binaries within 7.5" does appear to affect all planet classes, but the apparent radius distribution is more similar to the unblended distribution than that describing binaries within 45", suggesting that in the PLATO priority classes, contamination by blending within the same pixel is of limited importance.

In P1, for example, (Fig. 7.26, panels (e) and (f)), giant planets and those too large to be considered planets are dominated by the unblended population, while blending within the same pixel is more significant in terrestrial and Neptune-like planets. The minimum in the unblended population between binaries with a compact component and those without is more muted in blending within the same pixel, affecting observations of Neptune-like planets.

Fig. 7.27, panels (a) and (b), show that for the P2 population, small number statistics make it difficult to track trends across apparent radius, but above an apparent radius of 10 $R_{\oplus}$ all the binaries within 45" are unblended. This is consistent with the bright stars in the P2 population



(a) Unblended binaries, LLN (b) Unblended binaries, LLS

(c) Planets, LLN (d) Planets, LLS

(e) Blended binaries within 7.5", LLN (f) Blended binaries within 7.5", LLS

Fig. 7.24 Synthetic binaries, apparent planet radius against distance, P5, 24 cameras, compared with synthetic planets, P5, 24 cameras. A cut off is applied to the synthetic unblended eclipsing binary population to limit it to those binaries most likely to mimic planets.

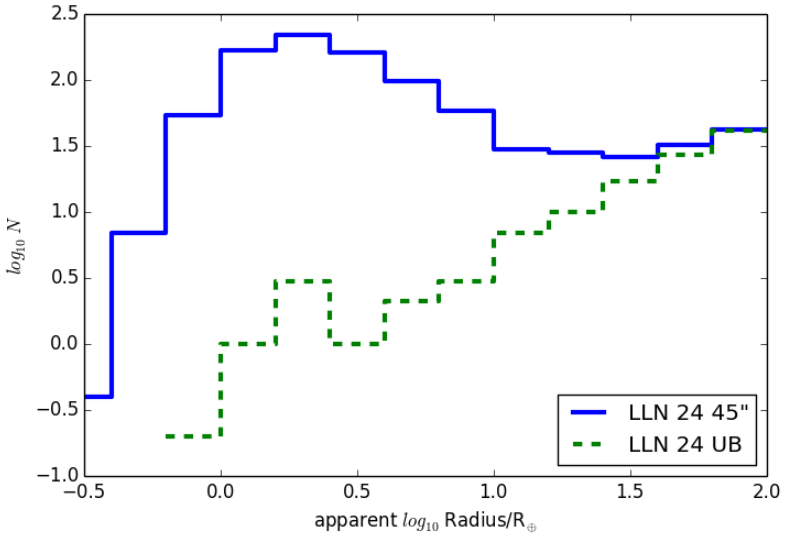


Fig. 7.25 Binaries within 45" of the target compared to unblended binaries, LLN P1, 24 cameras. Both samples have maxima at $0.2 < log\ R/R_\oplus < 0.4$. In the unblended sample this maximum is dominated by eclipsing binaries in which one component is a WD, while in the blended sample dilution by blending of distant binaries is responsible for the maximum.

($V \leq 8$) dominating their neighbourhood: background blends are likely to be unobservable, due to the size of the magnitude difference Δm (Eq.1.14). However, these are the binaries most likely to contaminate other targets, especially if only the wings of the PSF from this binary are observable in those other targets, emphasising the importance of a comprehensive catalogue to identify potential ephemeris matches.

The distribution of blended and unblended binaries is clearly different in the P4 population (Fig. 7.27, panels (c) and (d)). In the unblended population, the contribution of binaries with a compact object, dominating up to an apparent radius of $log\ R/R_\oplus$ 0.4, is distinct from the grazing binary population, which peaks at $1.2 < log\ R/R_\oplus < 1.6$. The blended population rises and falls smoothly in a way that appears to bear no relationship to the unblended population, especially as distance from the target increases. As previously noted, there is also an upper limit for the unblended population based on the maximum radius of an M-class star. Blended binaries appear to exceed this limit because δF is being assessed against the target, not against the true parameters of the blended eclipsing binary.

The P5 population contains the same class of stars as P1 (F5-K7), but the magnitude limit is fainter, $V < 13$ rather than $V < 11$. The effects of this can be seen in Fig. 7.27, panels (e) and (f): while the different contributions of compact binaries at smaller apparent radii and other binaries at larger apparent radii can still be distinguished in the unblended population,

there is a broad maximum in the blended population to 45" and this occurs at a larger apparent radius than the peak due to compact binaries in the unblended population. This is in contrast to the trends seen in Fig. 7.26, panels (e) and (f) (P1).

The plots illustrating the populations $V < 26$ and $V < 16$, F5-M, Fig. 7.26, panels (a) to (d), would indicate that this is part of a trend: as the magnitude limit becomes fainter, more contamination emerges of the Neptune-like planet population by false positives arising from blending with background binaries.

The actual level of potential contamination for each individual target will depend on its magnitude, which will define its equivalent of the *Kepler* exclusion radius within which a background blend cannot be separated from the PSF of the target. The actual degree of blending reflected in PLATO observations will probably lie between blends within the same pixel and blends within the same imagette.

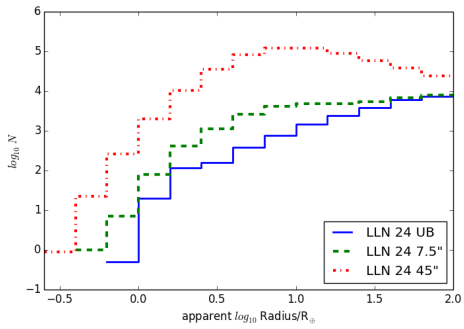
7.4.5 Further considerations

Ephemeris matches

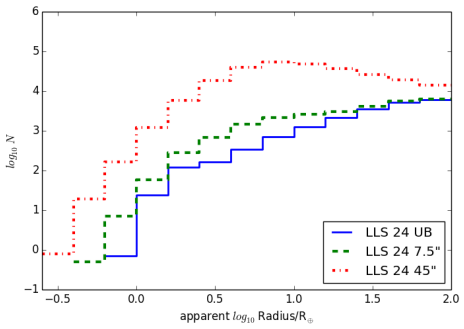
Background eclipsing binaries where only the wings of the PSF are observed, which should produce an ephemeris match in an appropriate catalogue, are more likely to mimic terrestrial planets. In the case of an ephemeris match to a neighbouring pixel, it may be possible through careful data analysis, and through consideration of the whole imagette, to eliminate such signals from further consideration. The PLATO Definition Study Report states that PLATO will undertaken quarterly slews, as did *Kepler*. With *Kepler*, one consequence of the quarterly slews was that it is easier to identify blends with background binaries, due to the relative change in position of the stars on the pixels. *Kepler* has smaller pixels than *PLATO*, but the same effect may well be identifiable in PLATO observations. Further ephemeris matches to the PSF wings of more distant binaries may arise from instrumental effects, as was the case through charge transfer along columns in *Kepler*. This problem could be mitigated if the cameras are not perfectly aligned, as a column in one camera may be slightly offset from a column in another camera. As indicated above, a small degree of misalignment is anticipated in the PLATO Definition Study Report. These cases have not been considered in this study.

White dwarfs

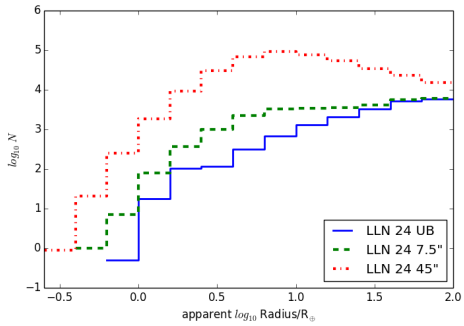
Compact objects simulated in this study and detectable by PLATO are exclusively white dwarfs rather than neutron stars and may mimic planets larger than their actual radius, because they are self luminous, so the drop in luminosity when they pass behind their companion is more noticeable than it would be for a planet, while the drop in luminosity when they pass in



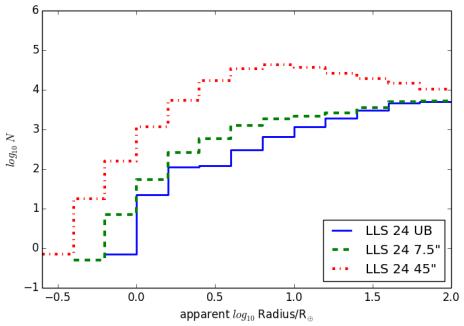
(a) $V \leq 26$, LLN



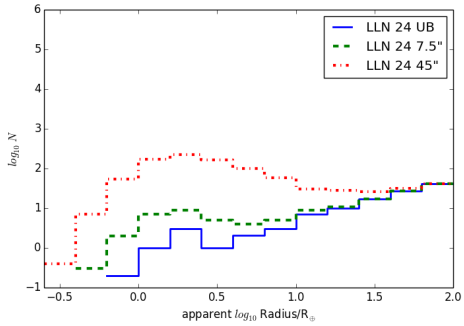
(b) $V \leq 26$, LLS



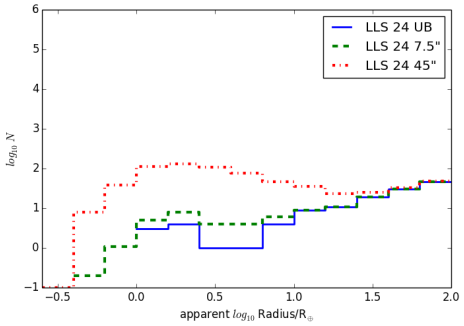
(c) $V < 16$, F5-M, LLN



(d) $V < 16$, F5-M, LLS

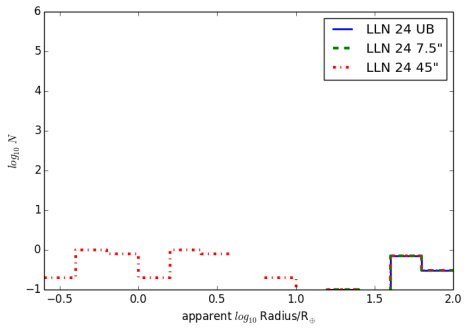


(e) P1, LLN

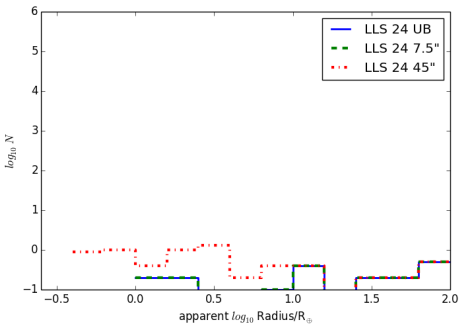


(f) P1, LLS

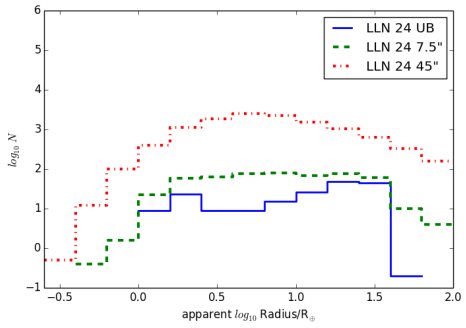
Fig. 7.26 Synthetic binaries within 45" and 7.5" of the target compared to unblended binaries, 24 cameras.



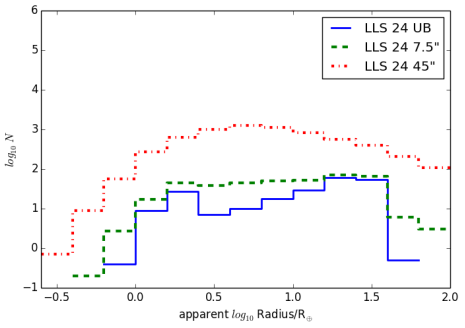
(a) P2, LLN



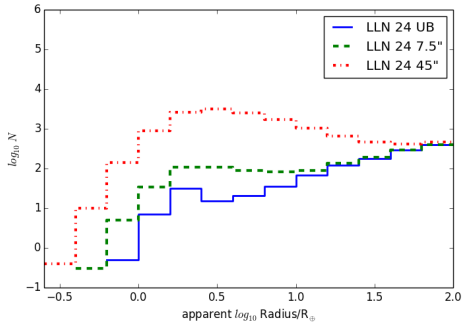
(b) P2, LLS



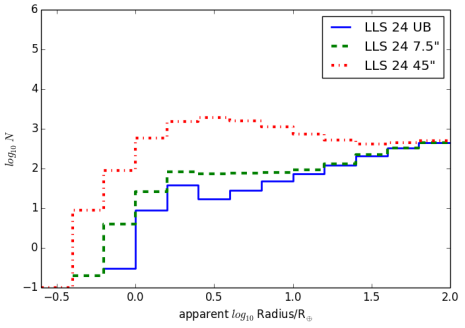
(c) P4, LLN



(d) P4, LLS



(e) P5, LLN



(f) P5, LLS

Fig. 7.27 Synthetic binaries within 45" and 7.5" of the target compared to unblended binaries, 24 cameras.

front of their companion is less noticeable because they cover only a small area of the primary and, of course, are continuing to contribute their own luminosity to the combined output.

Section 1.5.2 illustrates, with the case of KOI-256, the difficulties in confirming that a potential planet is in fact a white dwarf. Spectroscopic follow up by Muirhead et al. (2013) was required to confirm that the temperature of the WD at $T_{eff} \approx 7100 \pm 700$ K: *Kepler* analysis estimated the temperature of the "planet" at 734 K. As described in Section 1.5.2, this is because the *Kepler* analysis started with the assumption that the transiting body was an exoplanet, not a white dwarf. Examples such as KOI-256 emphasise that spectroscopic follow up is essential in confirming planets: it cannot be done from photometric light curves alone.

My analysis of synthetic eclipsing binary samples treats as the primary eclipse the deeper eclipse, whether it is the evolutionary primary passing in front of the evolutionary secondary or the other way around. If the drop in flux when a white dwarf is occluded is significant, but that when the white dwarf passes in front of its companion is not, especially in the case of a background blended binary, the equivalent of the "stellar eclipse" flag in the *Kepler* analysis may not be triggered, making it more likely that such systems would be forwarded as potential planets for further analysis. Because a WD is likely to be hotter than its companion, as was the case in KOI-256, where the MS companion has an estimated T_{eff} of 3450 ± 50 K, it is likely that when the WD is occluded, the eclipse depth will be deeper than when the WD occludes its companion.

7.5 Summary: Chapter 7

The work described in this chapter has set out to derive, not the true number of eclipsing binaries, rather the number of eclipsing binaries detectable in a transiting exoplanet survey. Using the *Kepler* Eclipsing Binary Catalogue, estimates of the distribution by mass ratio at $P < 10$ days have been made. The synthetic eclipsing binary population has then been weighted to resemble the observed population by period distribution.

These weightings are approximations and should be treated as such. They compensate for the flat initial period distribution in BiSEPS. They indicate a deficit in the synthetic eclipsing binary population at $10 < P/\text{days} < 40$, which bears further investigation.

The most significant limitation in my simulations is that the BiSEPS models are based on zero eccentricity, while especially at $P > 10$ days a wide variety of eccentricities are seen in nature. Eccentricity will affect transit duration, and incorporating eccentricity into `BiSEPS` is recommended for future work.

Estimates of the numbers of eclipsing binaries that will contaminate PLATO observations, at three levels of blending, are described in this chapter. The next chapter draws this together with the estimates of the number of transiting exoplanets observable by PLATO (Chapter 6)

to predict the effect of contamination by blended and unblended eclipsing binaries on PLATO observations.

Chapter 8

Interpretation of results

Using the BiSEPS population synthesis code (Chapter 2), light curves from JKTEBOP (Chapter 3), a synthetic Galaxy (Chapter 4) and the predicted precision of PLATO (Chapter 5), estimates of the numbers of exoplanets detectable by PLATO (Chapter 6) and the number of eclipsing binaries detectable by PLATO (Chapter 7) have been derived. Formal uncertainties have been derived in both the planet population (Appendix C) and in the eclipsing binary population (Appendix D). This chapter brings together the work described in the previous chapters to present ratios of planets to binaries.

8.1 Systematic uncertainties

8.1.1 Exoplanets

The estimates of the intrinsic exoplanet distribution have been derived from data from *Kepler* DR25, hosted on NExSci, as described in Chapter 6. However, as described in Section 6.2.1, the data on NExSci was derived on the basis that each *Kepler* target was a single star (Ciardi et al. (2015) and references therein) uncontaminated by blending with background stars. The stellar parameters were as derived from photometry and stellar evolutionary models for the *Kepler* Input Catalogue, unless updated by subsequent spectroscoopy. Work by Ciardi et al. (2015), Hirsch et al. (2017) and Furlan et al. (2017), described in more detail in Section 6.2.1, demonstrate that this sample is likely to be affected by dilution by blending with both bound and unbound companions not accounted for in *Kepler* DR25, so the estimated radii of the confirmed *Kepler* planets will, in a proportion of cases, be systematically too large or too small. Parallaxes and estimates of T_{eff} derived from spectroscopy from Gaia DR2 will also inform our understanding of the true stellar parameters of *Kepler* hosts.

The intrinsic exoplanet distribution 'zeta', based on the *Kepler* sensitivity for a single long cadence integration, requires a high normalisation factor of 15. The intrinsic exoplanet

Table 8.1 Comparison of *Kepler* sensitivity, 30 min rms photometric precision, *Kepler* Guest Observer Page (30 minute integration), 6.5 hr benchmark estimate, *Kepler* Instrument Handbook, and minimum rms CDPP, 6 hr as recorded on MAST for all stars of equivalent *Kp* (rms CDPP is assessed on a target-by-target basis).

Magnitude *Kp*	single long cadence (ppm)	benchmark estimate (ppm)	rms CDPP, 6 hr minimum (ppm)
12.0	80.5	22.7	29.246
12.5	107	28.9	24.022
13.0	147	38.1	29.649
13.5	206	52.1	39.155
14.0	296	73.7	46.089
14.5	436	107.3	61.815
15.0	656	160.2	82.848
15.5	1003	243.6	107.319
16.0	1552	375.7	173.525

distribution 'omega', based on the benchmark estimate in the *Kepler* Instrument Handbook, captures the transit depth against magnitude for most confirmed *Kepler* planets, but at this sensitivity cannot be complete in the observed population due to issues discussed in Section 6.2.3, especially affecting shallow transits. The derivation of the *Kepler* Multiple Event Statistic (MES), described in Section 5.1.2, subheading "Other considerations", makes it clear that phase folding multiple events increases the SNR, to the point where with a significant number of events, very shallow signals can be identified.

Kepler uses a Combined Differential Photometric Precision statistic (CDPP) to estimate white noise for each individual source (Christiansen et al., 2012). This is done on a quarter-by-quarter and target-by-target basis. The benchmark estimate from the *Kepler* Instrument Handbook is consistent with the 6 hr CDPP of main sequence stars as recorded on MAST. Table 8.1 compares the calculated eclipse depth, as shown in Table 6.1 with the minimum rms CDPP (6 hr) as recorded on MAST, for stars with equivalent *Kp*. This minimum value will correspond to the quietest main sequence stars: red giants have a higher rms CDPP.

Christiansen et al. (2012) demonstrated that, as integration time increases, the rms CDPP of a non-variable star falls, while in a variable star it increases (their Fig. 7 and Fig. 5 respectively).

Seader et al. (2015) and references therein summarise *Kepler* data processing as follows: pixel level data was cleaned and sent to the Photometric Analysis pipeline, to produce a flux time series for each target. Systematic variations were removed by the Pre-search Data Conditioning pipeline. Each corrected flux time series was then analysed by the Transiting Planet Search pipeline, searching for periodic reductions in flux. If certain criteria were

met, an event would be granted Threshold Crossing Event (TCE) status and further events, indicative of a multi-planet system, were searched for in the same flux time series. Further tests were then carried out before *Kepler* Object of Interest status was granted. For further information on each of these steps, see Seader et al. (2015) and references therein.

Data was acquired in 29.4 minute integrations, and these 'long cadence' integrations are the basis for all further analysis: see Section 6.1.1. The MES (Multiple Event Statistic) phase folds single events, thereby improving the signal to noise ratio as the square root of the number of events, when both signal and noise are reasonably constant (Tenenbaum et al., 2012). In effect, the red line in my Fig. 6.1 represents the average CDPP for a single event, while the black line represents the average CDPP for all transit events combined; and those planets between the red and black lines are only detectable when multiple events are considered (J. Coughlin, priv. comm, 30 May 2018).

I regard 'zeta' and 'omega' as representing limiting cases that bracket the true intrinsic distribution of confirmed *Kepler* planets. 'Zeta' underestimates the sample as a single 30 minute integration, on which 'zeta' is based, misses a significant number of shallow transits around faint stars which the *Kepler* pipeline can identify through phase folding. 'Omega' overestimates the sample as it is not complete.

8.1.2 Eclipsing binaries

As described in Chapter 7, the unblended eclipsing binary population was calibrated by initial mass ratio and by period, to resemble that which is detectable in a transiting exoplanet survey.

The calibration by initial mass ratio of binaries with $P < 10$ days, which indicates that binaries with a high mass ratio are common at short periods, is consistent with the findings from Raghavan et al. (2010) that inner pairs of triple systems are more likely to consist of equal mass components.

The calibration by period is much more prone to systematic errors. Table 7.7 and Fig. 7.19 describe the match that can be achieved as a result of the calibration, and Table 7.7 and Fig. 7.20 report the coefficients used in the period calibration. The period calibration also depends on the use of the initial mass ratio calibration. The period calibration is less well defined and carries a larger uncertainty.

8.2 Ratio of planets to binaries in LLN and LLS

My results are presented in Fig. 8.2 to Fig. 8.7, which illustrate the ratio of planets to binaries with the following parameters: true planet radius and apparent planet radius allowing for blending within the pixel; planet distributions based on both the 'zeta' and 'omega' PDFs; and

blending with background binaries to distances of 0.0" (unblended), 7.5" (within the same pixel) and 45.0" (within the same imagette). The PLATO precision is based on a one hour integration in both the synthetic eclipsing binary and the synthetic exoplanet populations.

Planet to binary ratios where planets were simulated using the 'zeta' and 'omega' PDFs are distinguished in the figures by the use of the labels ζ and ω on the right hand side.

The colour scheme used in these figures and described in Table 8.2 indicates the probability that an event is caused by a planet or a binary, although it should be noted that other sources of false positives and false alarms, such as variable stars and instrumental effects, are not taken into account.

Formal uncertainties are not indicated in these plots, because these are generally smaller than the difference between the numbers generated by the two intrinsic planet distributions, most pronounced among terrestrial planets. Full tables with the planet to binary ratios, including the fractional uncertainties, $\delta X/X$, are included in Appendix E, Table E.1 to Table E.12.

It should also be noted that binaries will be detected in radius bins indicated as having a high confidence of returning a planet, and planets will be detected in bins indicated as having a high confidence of returning a binary.

Each of these six figures is divided into an upper panel where the true radius of the planet is considered, and a lower panel where blending within the same pixel is taken into account, which may place a given planet in a different radius bin to that indicated by its true radius. Blending within the same pixel is likely to be the condition under which PLATO observations are made.

Fig. 8.1 presents the data in an alternative fashion for the most significant result: P5, LLS, considering blending with binaries within 7.5" of the target. P5 is the most significant population as there are the numbers of both planets and eclipsing binaries to make a meaningful assessment possible. LLS is the more significant field in this context as the number of planets is predicted to be higher and the number of background eclisping binaries is expected to be lower than in LLN, so any regions in the parameter space where fewer planets are predicted than binaries should be cause for concern. Blending within 7.5" is the most significant blending radius as blending on the same pixel will be harder to distinguish than blending across two or more pixels. The uncertainties are derived by considering the formal uncertainties in both the planet and binary populations, and are large. Four panels are also required for each population in each field at each blending radius: two for each PDF, with one panel for each PDF considering the true (unblended) radius and another considering the apparent radius after blending, which is what will be observed.

The conclusions I would draw from this figure are: (1) the smallest planets may appear to outnumber binaries, but the error bars are so large that no reliable conclusions can be drawn;

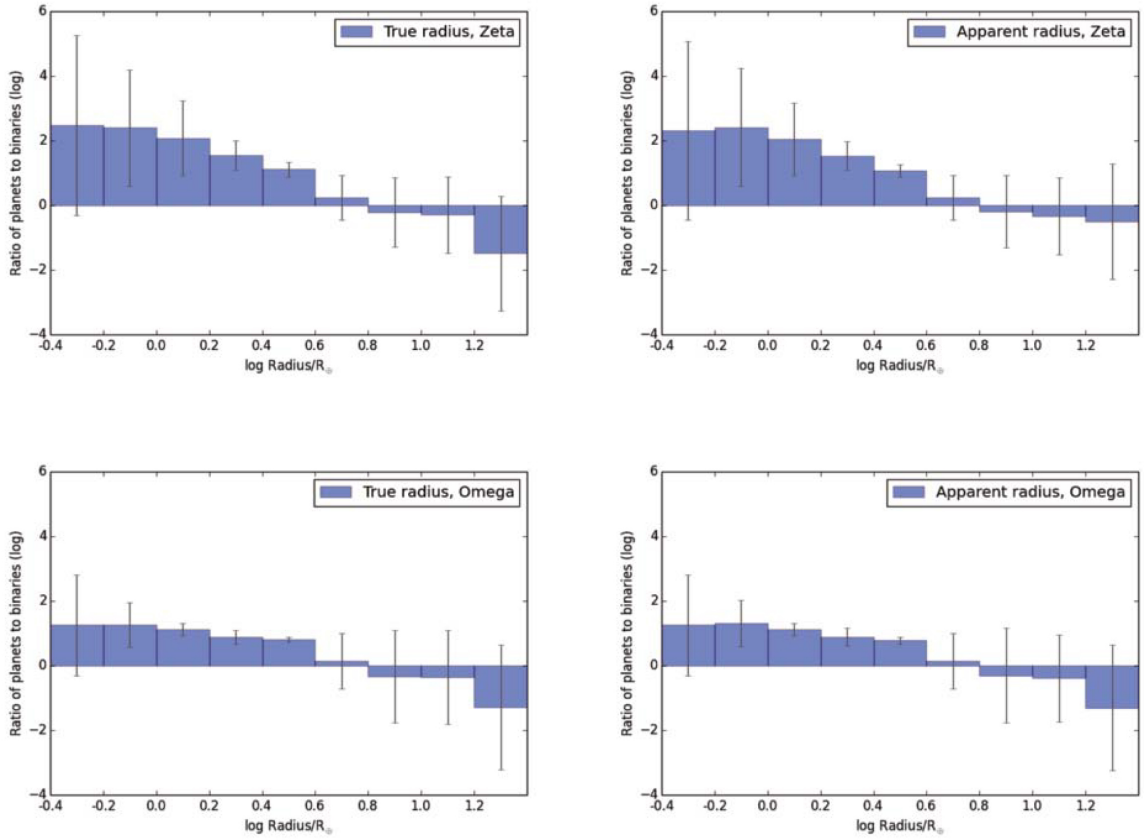


Fig. 8.1 Ratio of planets to blended binaries, LLS, P5, as a bar chart, blending with eclipsing binaries within 7.5". Top left: true radius, planet distribution based on 'zeta'. Top right: apparent radius, planet distribution based on 'zeta'. Bottom left: true radius, planet distribution based on 'omega'. Bottom right: apparent radius, planet distribution based on 'omega'.

(2) giant planets, which the PLATO consortium has already indicated will be treated as false positives (Rauer et al., 2014), are likely to be outnumbered by contaminating binaries; and (3) the observed ratio of planets to binaries will depend critically on whether ζ or Ω is a truer reflection of the intrinsic exoplanet distribution.

In the light of the very large uncertainties, I consider it important to examine the big picture in Fig. 8.2 to Fig. 8.7 to get a sense of which populations will be most affected by blending with background eclipsing binaries. I would encourage those using Fig. 8.2 to Fig. 8.7 to concentrate on the colours rather the actual numbers presented, for a sense of where there may be problems in detecting planets.

Fig. 8.2 and Fig. 8.3 consider the ratio of planets to unblended eclipsing binaries. Fig. 8.4 and Fig. 8.5 consider the ratio of planets to binaries within 7.5" of a target star. Fig. 8.6 and Fig. 8.7 consider the ratio of planets to binaries within 45" of the target star. See the tables in Appendix C for numbers of planets and the tables in Appendix D for numbers of binaries, with random uncertainties stated in both sets of tables.

Fig. 8.2 to Fig. 8.7 are organised with the class most susceptible to blending (P4) at the top, and the class least susceptible to blending (P2) at the bottom. P2 is a subset of P1, P1 is a subset of P5 and P5 (with P4) is a subset of $V < 16$, F5-M.

Since in Fig. 8.2, Fig. 8.4 and Fig. 8.6, describing LLN, the same numbers of planets (unblended and blended) are used, as expected the ratio of planets to binaries decreases as blends further from the target are considered. The same applies to Fig. 8.3, Fig. 8.5 and Fig. 8.7, describing LLS. Even when the most distant blended eclipsing binaries are considered, however, the number of terrestrial planets ($-0.4 < log\ R/\mathrm{R}_\oplus < 0.2$) is greater than the number of eclipsing binaries of similar apparent radius when the planets are simulated using the 'zeta' PDF. When the planets are simulated using the 'omega' PDF and blending within 45" is considered, the number of binaries outnumbers the number of planets. Giant planets ($log\ R/\mathrm{R}_\oplus > 0.8$) tend to have similar numbers to binaries, or be outnumbered by binaries, however fine the angular resolution that can be achieved, and the ratios tend to be similar whether the planet distribution was derived with 'omega' or 'zeta'.

The results in Fig. 8.2 to Fig. 8.7 are as expected from Chapter 6 and Chapter 7: with twice as many binaries in the same imagette in LLN than LLS, but more terrestrial planets in LLS than LLN, it would appear that the most fruitful place to search for terrestrial planets, when blending is taken into account, would be in LLS (Fig. 8.6 and Fig. 8.7). If blended and unblended binaries could be separated, so that only unblended binaries need to be taken into account, LLN becomes more fruitful (Fig. 8.2 and Fig. 8.3): however, in practice this is very hard to achieve.

Phase folding in the PLATO pipeline may well allow signals from transiting planets and eclipsing binaries below the sensitivity of a one hour integration to be detected, in the same

Table 8.2 Guide to colours used in planet vs binary plots (Fig. 8.2 to Fig. 8.7).

Colour	Ratio planets: false positivies	Interpretation
Dark Green	> 10	Very high confidence this is a planet
Light Green	5-10	High confidence this is a planet
Yellow	2-5	Reasonable confidence this is a planet
Orange	1-2	Could be either a planet or a binary
Red	0.1-1	Low confidence this is a planet
Grey	0.01-0.1	Very low confidence this is a planet
Black	0.001-0.01	Extremely low confidence this is a planet
White	n/a	No planets (NP), not possible to assess or no binaries (NB), or both (NB NP)

way that the *Kepler* pipeline was able to recover transits of planets below the sensitivity of a single long cadence observation (Section 8.1.1). Tests on the population derived with 'omega' using a longer PLATO integration did allow 'detection' of significantly more terrestrial planets than was the case with a one hour integration. However, it is assumed that such phase folding would enhance the detectability both of terrestrial planets and of eclipsing binaries mimicking terrestrial planets, to a similar degree, and so the ratios of planets:binaries by radius bin presented here would not be affected.

The conclusion that LLS may have fewer background contaminants than LLN does depend on assumptions for the composition of the Galaxy. Our model uses a two-disc double exponential model, as described in Chapter 2, allowing for Drimmel extinction (Drimmel et al., 2003).

The spatial resolution across the celestial sphere in the Drimmel model for extinction is the same as the FIR COBE/DIRBE data; the NASA website describing the COBE/DIRBE data[1] states that the beam solid angles were computed based on the beam response over an area of $\pm$ 42.2 arcmin in scan and $\pm$ 36.2 arcmin cross-scan. A finer resolution would be beneficial in more closely matching the observable details of the Galaxy. Extinction in the third dimension, distance, also needs to be considered and Drimmel et al. (2003) describes how, through the use of NIR colour-magnitude diagrams in the Galactic plane, extinction can be derived up to a distance of 8 kpc.

As data from Gaia becomes available, with DR2 released in April 2018 and DR3 due in 2019, knowledge of the solar neighbourhood, both in terms of stellar density and molecular cloud distribution, will increase and it will be possible to further calibrate our models. As shown in Fig. 2.3, in the *Kepler* field, for which it is possible to compare Gaia DR1 with

[1]https://lambda.gsfc.nasa.gov/product/cobe/dirbe_beam_char.cfm

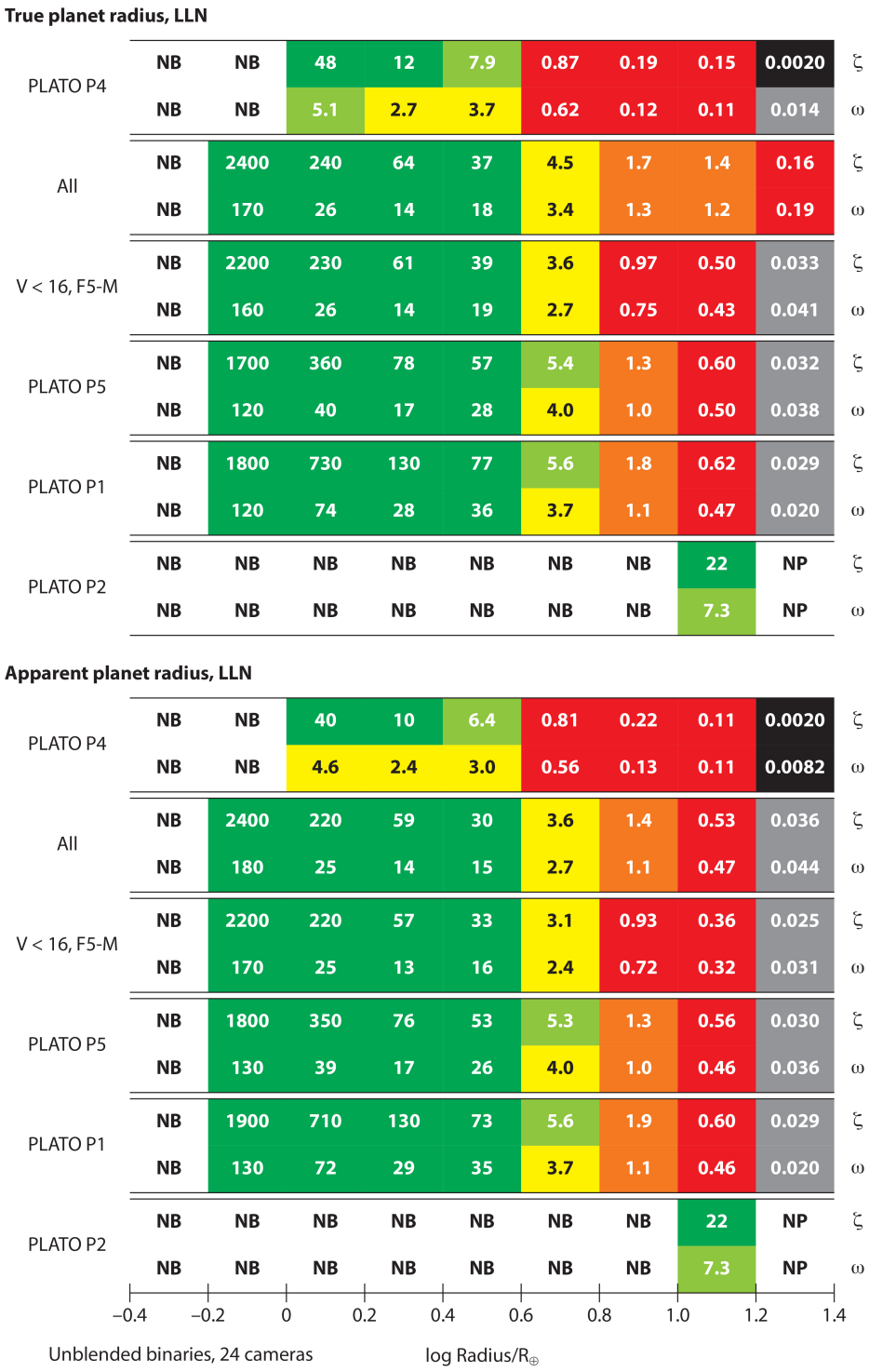


Fig. 8.2 Ratio of planets (true radius and apparent radius after blending within the same pixel) to unblended binaries, LLN. See Table 8.2 for the key to the colours. ζ indicates a synthetic planet population derived with the 'zeta' PDF, ω indicates a synthetic planet population derived with the 'omega' PDF. See Table E.1 and Table E.2 for formal uncertainties $\delta X/X$.

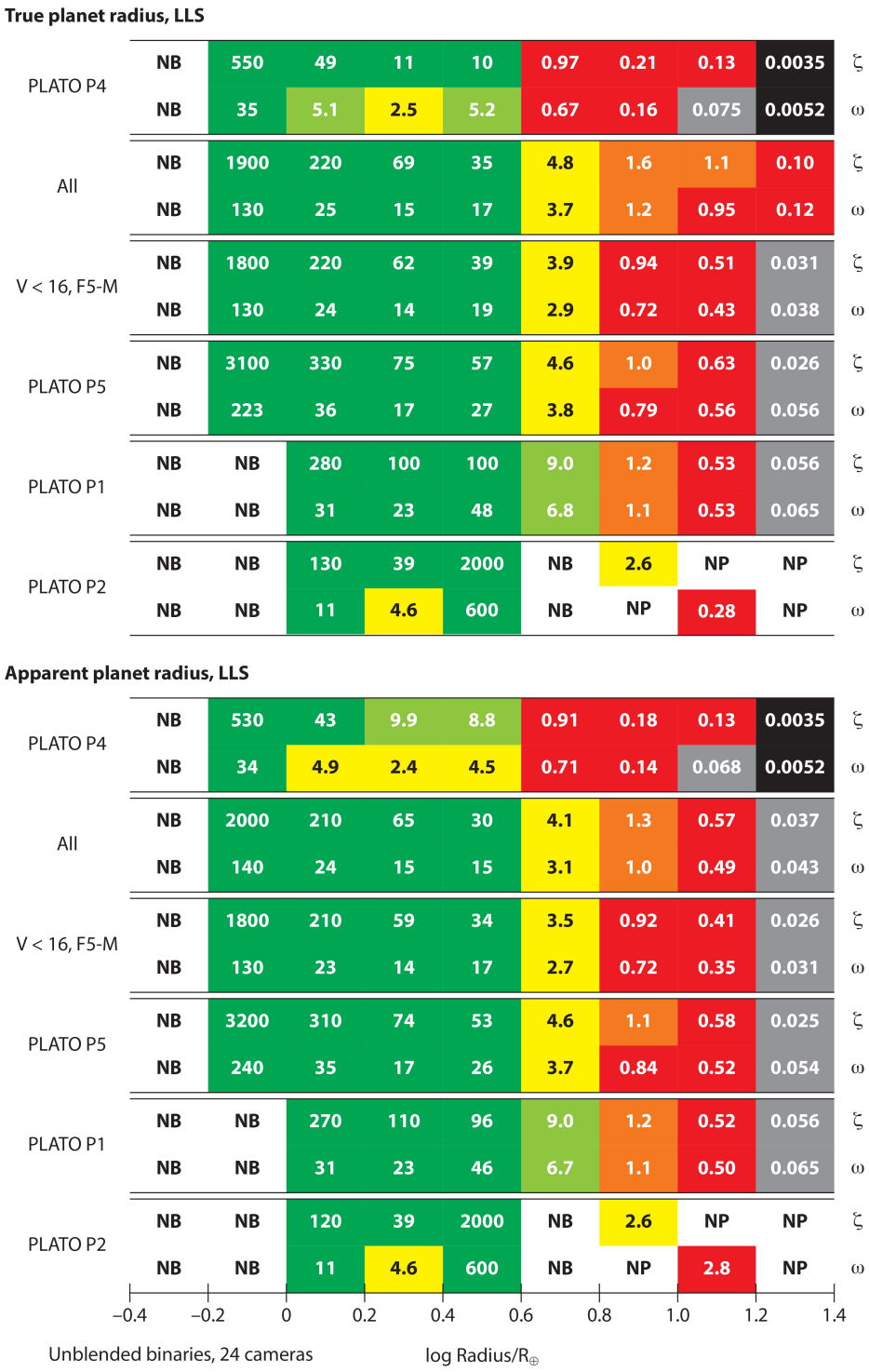


Fig. 8.3 Ratio of planets (true radius and apparent radius after blending within the same pixel) to unblended binaries, LLS. See Table 8.2 for the key to the colours. ζ indicates a synthetic planet population derived with the 'zeta' PDF, ω indicates a synthetic planet population derived with the 'omega' PDF. See Table E.3 and Table E.4 for formal uncertainties $\delta X/X$.

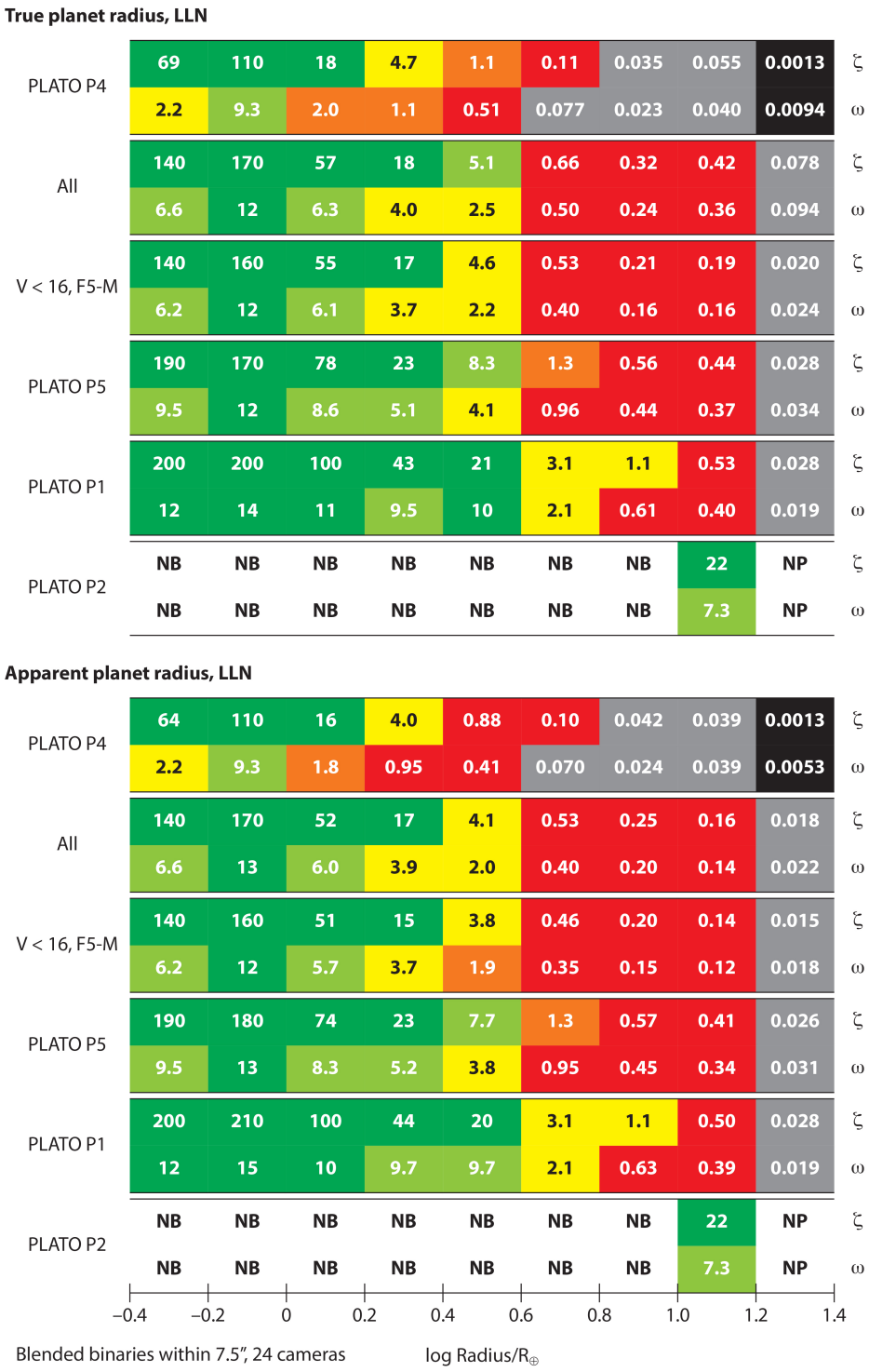


Fig. 8.4 Ratio of planets (true radius and apparent radius after blending within the same pixel) to binaries within 7.5", LLN. See Table 8.2 for the key to the colours. ζ indicates a synthetic planet population derived with the 'zeta' PDF, ω indicates a synthetic planet population derived with the 'omega' PDF. See Table E.5 and Table E.6 for formal uncertainties $\delta X/X$.

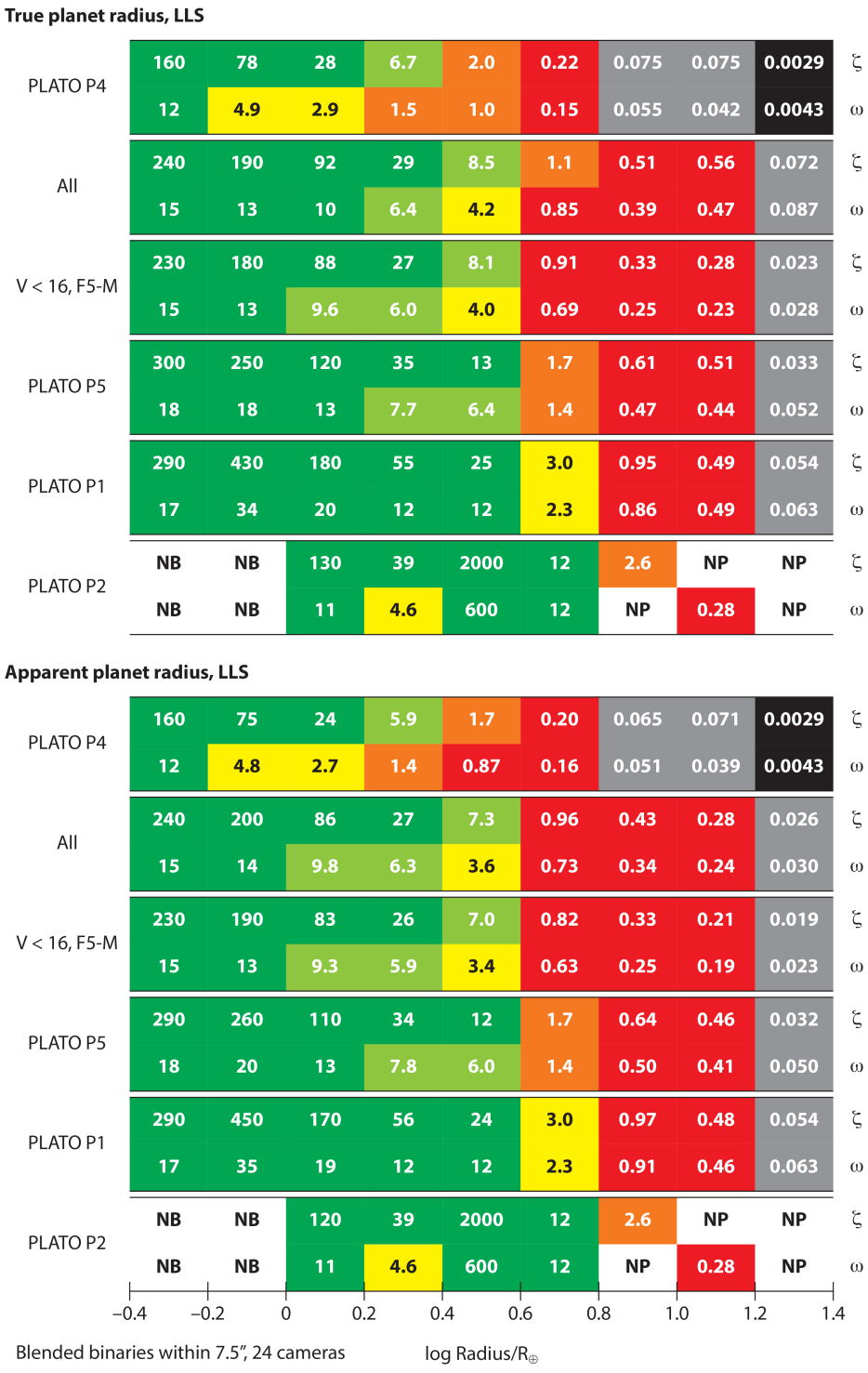


Fig. 8.5 Ratio of planets (true radius and apparent radius after blending within the same pixel) to binaries within 7.5", LLS. See Table 8.2 for the key to the colours. ζ indicates a synthetic planet population derived with the 'zeta' PDF, ω indicates a synthetic planet population derived with the 'omega' PDF. See Table E.7 and Table E.8 for formal uncertainties $\delta X/X$.

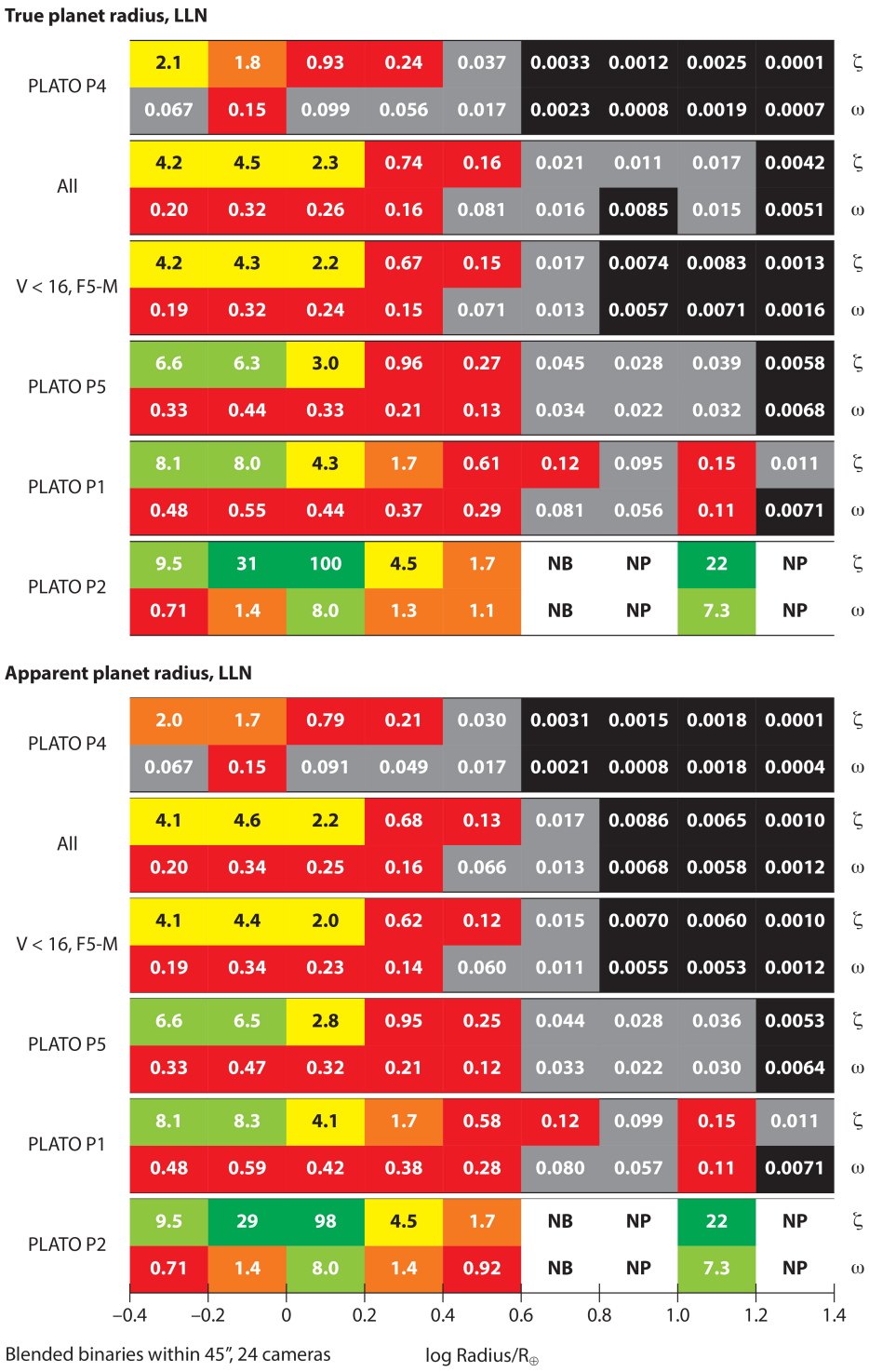


Fig. 8.6 Ratio of planets (true radius and apparent radius after blending within the same pixel) to binaries within 45", LLN. See Table 8.2 for the key to the colours. ζ indicates a synthetic planet population derived with the 'zeta' PDF, ω indicates a synthetic planet population derived with the 'omega' PDF. See Table E.9 and Table E.10 for formal uncertainties $\delta X/X$.

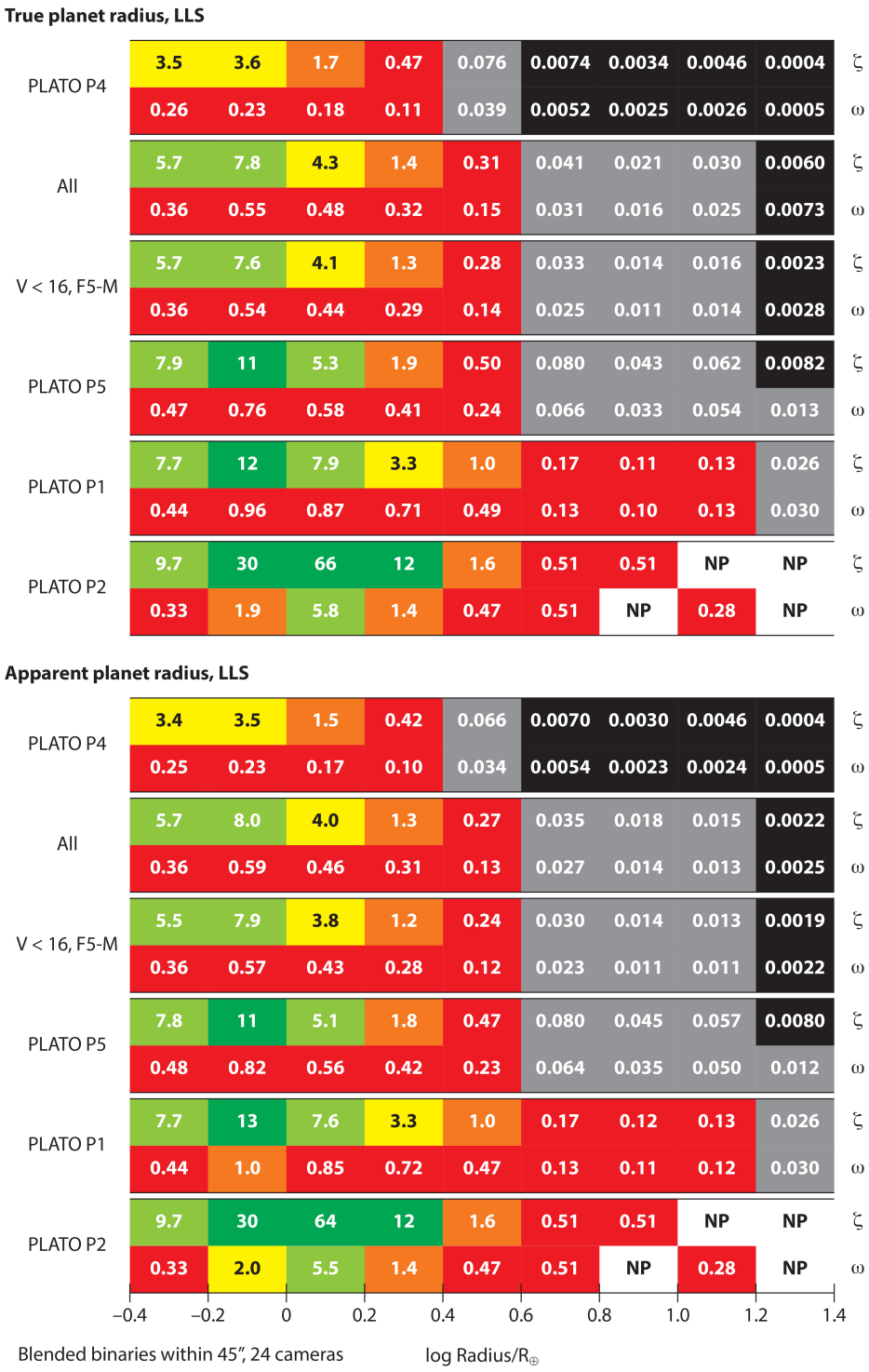


Fig. 8.7 Ratio of planets (true radius and apparent radius after blending within the same pixel) to binaries within 45", LLS. See Table 8.2 for the key to the colours. ζ indicates a synthetic planet population derived with the 'zeta' PDF, ω indicates a synthetic planet population derived with the 'omega' PDF. See Table E.1 and Table E.2 for formal uncertainties $\delta X/X$.

the *Kepler* Input Catalogue on MAST, Gaia DR1 includes about 60% of the known stellar population. Because of the incompleteness of Gaia DR1 compared to MAST, Gaia DR1 could not be used to reliably calibrate other regions in the synthetic fields.

My results demonstrate that possible giant planets are likely to be overwhelmingly eclipsing binaries, while conclusions drawn about terrestrial and Neptune-like planets are sensitive to the intrinsic exoplanet distribution and the degree of blending. M-class stars are at greater risk of contamination from background eclipsing binaries than other classes.

Fig. 8.2 to Fig. 8.7 cover ratios utilising 24 cameras, the number of cameras in the design at the time of writing. Planet to binary ratios utilising 28 and 32 cameras are very similar to those utilising 24 cameras, and differences due to uncertainty in the models (random and systematic) are more significant than the difference between the ratios in different numbers of cameras. This leads to the conclusion that, while reducing the number of cameras to 24 will reduce the number of planets that can be observed, especially in P4 and P5, it will not significantly change the probability that a given event is caused by a planet or a binary.

8.2.1 Summary by PLATO priority population

Table 8.3 summarises for terrestrial ($-0.4 < log\ R/R_\oplus < 0.2$), Neptune-like ($0.2 < log\ R/R_\oplus < 0.8$) and giant planets ($0.8 < log\ R/R_\oplus < 1.4$), whether more planets or more binaries will be detected in a given field with a given degree of blending. It should be used in conjunction with Fig. 8.2 to Fig. 8.7.

In P1, within the same imagette, there are more Earth-like and “Small Neptune”-like planets:binaries in LLS (Fig. 8.7) than in LLN (Fig. 8.6), whichever PDF is used to simulate the planets. For Mars-like, “Large Neptune”-like and giant planets the ratios are similar in the two fields, given the uncertainties. Within the same pixel, ratios for terrestrial planets and the two Small Neptune classes are more favourable in LLS (Fig. 8.5), for Large Neptunes the ratio is the same in both fields and for giant planets the ratios are more favourable in LLN (Fig. 8.4) when the planet population is derived with ‘zeta’, although in both fields what is believed to be a giant planet is more likely to be a binary. When the planet population is derived with ‘omega’, the results concerning giant planets and blending within the same pixel are more mixed. If only unblended binaries are included, in some radius bins the ratio of planets:binaries is higher in LLN (Fig. 8.2), in others it is higher than in LLS (Fig. 8.3).

In P1, terrestrial planets will outnumber binaries in both fields, at all levels of blending where the planet population is derived using ‘zeta’, while binaries mimicking giant planets will outnumber true giant planets, again at all levels of blending and whichever PDF is used. Where the planet population is derived using ‘omega’, binaries mimicking terrestrial planets are likely to outnumber true terrestrial planets when blending within the same imagette is

considered. The detection of terrestrial planets is the priority for PLATO, and it is stated in Rauer et al. (2014) that all potential giant planets will be treated as false positives. Given that the likely level of blending will be between same pixel and same imagette levels, the evidence from Fig. 8.4 to Fig. 8.7 suggests that of the two proposed observing strategies (Section 1.3), 3+1 years, with LLS being observed for three years and LLN for one, would be most beneficial in identifying Earth-like planets.

In P2, small number statistics, especially for the binary population, make analysis of this priority class uncertain, whatever the level of blending. Stars in this class are more likely to contaminate other targets of interest than to be contaminated themselves. No conclusions are drawn for this class of essentially rare stars in this work.

In P4, within the same imagette, while the ratios for LLS look a little more promising than those in LLN, if there is a risk of blending out to 45" in a PLATO observation, any planet with an M-class host should be treated with extreme caution. This caution is especially indicated in populations derived using 'omega'. Within the same pixel, events that appear to be caused by terrestrial planets orbiting M-class stars are likely to indicate the presence of genuine planets, but in both fields these may actually be larger planets affected by dilution by blending with other stars within the pixel. If unblended binaries only are considered there is little significant difference between the two fields.

Anything other than a terrestrial planet should be treated with caution in either field, both because our results indicate more binaries than planets at most degrees of blending and because giant planets with M-class hosts are expected to be essentially rare (Fressin et al. (2013) and references therein).

In P5, the ratios indicate that, within the same imagette, the planet to binary ratio for terrestrial planets is sensitive to the intrinsic exoplanet PDF. Populations derived with 'zeta' indicate more terrestrial planets than binaries, while populations derived with 'omega' indicate more binaries than terrestrial planets. Within the same imagette, Neptune-like planets should be treated with caution, and giant planets with extreme caution. Within the same pixel, at all planet radii the ratio of planets to binaries is higher in LLS than LLN, although as planet radius increases the distinction becomes less clear. When unblended binaries only are considered, the ratios of terrestrial planets:binaries are higher in LLS than LLN in some bins, and higher in LLN than LLS in other bins.

8.2.2 Stellar density along the line of sight

The trends indicating that terrestrial planets are more likely to be identified in LLS than LLN, despite the fact that LLS is further from the Galactic centre so should have a less dense stellar

Table 8.3 Broad classification of planet to binary ratios. Rocky: $-0.4 < log\ R/R_\oplus < 0.2$. Neptune: $0.2 < log\ R/R_\oplus < 0.8$. Giant: $0.8 < log\ R/R_\oplus < 1.4$. LLN is Long Look North field, LLS is Long Look South field. "Same imagette", "Same pixel" and "Unblended" refer to the distance at which blending with binaries is considered. Each broad grouping includes three subgroups. "More planets" indicates all boxes include more planets than binaries, "similar numbers" indicates some boxes have more planets than binaries and some have more binaries than planets, and "more binaries" indicates all boxes include more binaries than planets.

Priority Population	LLN Rocky	LLN Neptune	LLN Giant	LLS Rocky	LLS Neptune	LLS Giant
Same Imagette						
P1	Depends on PDF	Similar numbers	More binaries	Depends on PDF	Similar numbers	More binaries
P2	Depends on PDF	More planets	Very few planets	Depends on PDF	Similar numbers	Very few planets
P4	Depends on PDF	More binaries	More binaries	Depends on PDF	More binaries	More binaries
P5	Depends on PDF	More binaries	More binaries	Depends on PDF	Similar numbers	More binaries
Same Pixel						
P1	More planets	More planets	Similar numbers	More planets	More planets	More binaries
P2	No binaries	No binaries	Very few planets	More planets	More planets	Very few planets
P4	More planets	Similar numbers	More binaries	More planets	Similar numbers	More binaries
P5	More planets	More planets	More binaries	More planets	More planets	More binaries
Unblended						
P1	More planets	More planets	Similar numbers	More planets	More planets	Similar numbers
P2	No binaries	No binaries	Very few planets	More planets	More planets	Very few planets
P4	More planets	Similar numbers	More binaries	More planets	Similar numbers	More binaries
P5	More planets	More planets	Similar numbers	More planets	More planets	Similar numbers

population (confirmed by numbers of stars at $V < 16$ and $V < 26$ in both synthetic fields), may be explained by assuming the following:

(a) PLATO prioritises bright, nearby stars.

(b) To observe the LLS field, it is necessary to look through the Galactic plane. This is not the case to observe the LLN field.

(c) `BiSEPS` uses a double exponential to calculate the stellar density of both the thin and thick discs (Section 2.1.2). This is consistent with observational evidence (Section 1.6) (Jurić et al., 2008). Fig. 8.8 compares the integrated stellar density along lines of sight at the central values of l for each field (65° LLN, 253° LLS), by $|b|$ and distance in kpc, before allowing for extinction. A line indicates the region in which the integrated stellar density is the same in each field as a function of $|b|$ and distance, before allowing for extinction. To the left of and above the line, the integrated stellar density along the line of sight is higher in LLS: to the right of and below the line, the integrated stellar density along the line of sight is higher in LLN. Note that at $|b| \gtrapprox 48.5°$, integrated stellar density along the line of sight is always higher in LLS, before allowing for extinction. Fig. 8.9 relates the lines of sight in Table 8.4 and Fig. 8.8 to the scale length used by the thin and thick disc.

(d) Hosts of terrestrial planets capable of being observed by PLATO are unlikely to be more than one scale height above or below the Galactic plane (Fig. 6.18), and therefore are likely to be more common in LLS. Fig. 6.16 and Fig. 6.17 indicate that stars in the P5 population capable of hosting planets detectable by PLATO with 24 cameras in a one hour integration are all at < 1.2 kpc. In Fig. 8.8 all stars at < 1.2 kpc are in the region where the integrated stellar density along the line of sight (before considering extinction) is higher in LLS than LLN, at all values of $|b|$.

(e) Binaries mimicking terrestrial planets in a one hour integration are likely to be highly blended (Fig. 7.24) and can be up to $\approx$ 10 kpc from Earth, and therefore, from Fig. 8.8, are likely to be far more common in LLN, especially at low values of $|b|$.

Evidence to support this explanation includes the fact that synthetic unblended binaries are more significant in LLS, where again the population at or close to the Galactic plane would be likely to be greater than in LLN. However, this conclusion is based on a synthetic Galaxy, not an observed one, and while Drimmel extinction has been included in our simulations, there may be significant variations between simulations and observations. The Galaxy is not isotropic, as the simulations assume. Star formation occurs in clusters as molecular clouds collapse. Over time, the clusters disperse due to dynamic forces and stars move independently around the Galactic centre.

Extinction explains the data in Table E.13 to Table E.30 (Appendix E), which compares synthetic fields 1 deg^2 in LLN and LLS positioned similarly in relation to the centre of each field. In some cases, where Fig. 8.8 and Table 8.4 lead to the expectation that, based on the

Table 8.4 Ratio of integrated stellar density along the line of sight, LLN/LLS, at $|b| = 30°$, by l and d, before considering extinction.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0.939	0.939	0.939	0.939	0.938
0.10	0.902	0.903	0.903	0.902	0.901
0.20	0.892	0.892	0.893	0.892	0.890
0.30	0.893	0.895	0.895	0.894	0.892
0.40	0.898	0.900	0.901	0.899	0.897
0.50	0.904	0.907	0.907	0.906	0.903
0.60	0.911	0.913	0.914	0.913	0.909
0.70	0.917	0.920	0.921	0.920	0.916
0.80	0.923	0.927	0.928	0.927	0.923
0.90	0.929	0.933	0.935	0.934	0.929
1.00	0.934	0.939	0.941	0.940	0.935
2.00	0.978	0.989	0.995	0.995	0.990
3.00	1.006	1.021	1.031	1.035	1.031
4.00	1.023	1.042	1.056	1.064	1.064
5.00	1.033	1.055	1.073	1.085	1.089
6.00	1.040	1.064	1.084	1.100	1.107
7.00	1.044	1.069	1.092	1.109	1.120
8.00	1.046	1.072	1.096	1.115	1.128
9.00	1.047	1.074	1.098	1.119	1.134
10.00	1.048	1.075	1.100	1.121	1.137

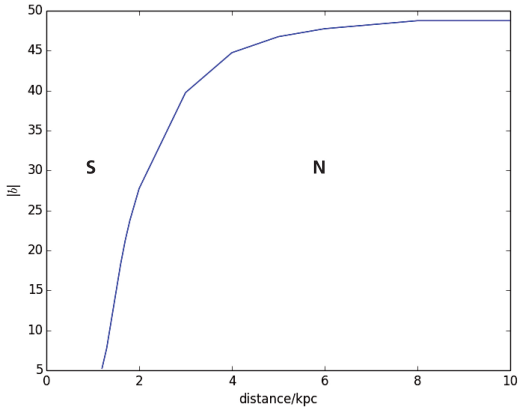


Fig. 8.8 Comparison of integrated stellar density along the line of sight, $l = 65°$, Northern hemisphere (LLN) and $l = 253°$, Southern hemisphere (LLS), before considering extinction. The line marks the region of equal integrated stellar density along the line of sight, by distance and $|b|$, in the two fields. In the region to the left of and above the line, integrated stellar density along the line of sight is higher in LLS. In the region to the right of and below the line, integrated stellar density along the line of sight is higher in LLN. The line crosses the x axis at 1.2 kpc. The curve is flat from 8 to 20 kpc.

double exponential alone, the integrated stellar density along the line of sight would be higher in either LLN or LLS, the opposite is true. Extinction is accounted for in this study: any intrinsic clumpiness of the local stellar density is not. And the evidence from Chapter 6 and Chapter 7 indicates that, whatever the local variation, over the full field the integrated stellar density at distances relevant to the detection of planets and unblended eclipsing binaries is higher in LLS, and at distances relevant to the detection of background eclipsing binaries is higher in LLN, in a one hour integration.

Observations are essential in determining the true stellar density along a given line of sight and, consequently, further analysis with Gaia DR2 and later Gaia releases is essential in confirming or refuting the conclusions presented here.

8.3 Summary: Chapter 8

This chapter has drawn together the work in all the proceeding chapters to estimate the ratio of planets to eclipsing binaries (blended and unblended) that PLATO will observe. Other types of contamination, such as instrumental effects and stellar variability, are not considered here. Therefore, the estimates presented are a ‘best case’ scenario.

When considering the ratio of terrestrial planets to binaries, the systematic uncertainties arising from the differences between the two intrinsic exoplanet distributions are greater than the random uncertainties in a given model.

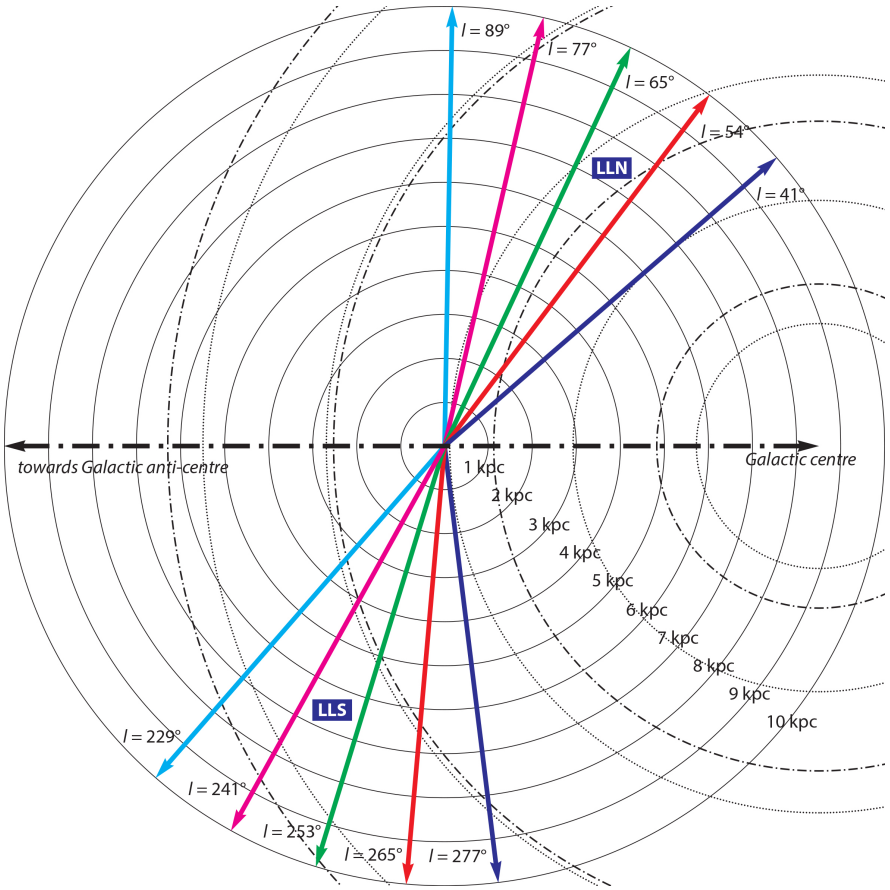


Fig. 8.9 Diagram to accompany Table 8.4. Solid circles indicate distance from Earth in kpc, arrows indicate lines of l and are coloured to indicate which lines are paired in Table 8.4. The dash-dot arrow indicates the line to (and from) the Galactic centre, stopping on the right hand side of the diagram at the Galactic centre. Dotted circles indicate radial scale lengths of the thin disc, dash-dot circles indicate scale heights of the thick disc. From Fig. 1.9, the bulge will only affect the region from $l = 330^\circ$ to $l = 30^\circ$, so is not relevant when considering the Long Look fields. In our simulations, in the PLATO priority populations, terrestrial planets are within 1.2 kpc of Earth, and all planets are within 2 kpc. At these distances, the radial scale length is less significant than the scale height. Background eclipsing binaries can be detected up to 10 kpc from Earth, and here the radial scale length is much more significant.

Taking Fig. 8.1 as a case in point, the smallest planets appear to have an impressive ratio of planets to binaries, but the uncertainties are large. Possible giant planets are more likely to be false positives. Estimates of the radius of planets is likely to be affected by dilution by blending. And estimates of terrestrial and Neptune-like planets:binaries based on Ω are half an order of magnitude lower than those based on ζ.

When blending with background binaries is included, whether on the pixel scale or the imagette scale, contamination is greater in the Northern field than in the Southern field. This is due both to the volume limited nature of stars in the PLATO priority populations, and to the deeper background population present in the Northern field. Data from Gaia DR2 and later Gaia data releases will be essential in assessing the reliability of this conclusion.

The final chapter in this thesis will present my conclusions, and will also draw together my recommendations for future work.

Chapter 9

Conclusions and Future Work

9.1 Conclusions

9.1.1 Background and method

PLATO, due to launch in 2026, is an ESA planet hunting mission which will survey large areas of sky (2250 deg^2) for long periods of time. The high levels of precision that can be obtained from the four groups of cameras, with six cameras per group, is offset by the increased risk of false positives due to the large pixel size (15.0"). This study aims to quantify the contribution of false positives from blended and unblended eclipsing binaries, and to compare this to the numbers of planets expected to be detected.

A further aim is to understand mechanisms of the formation of binaries and higher order multiples. Close binaries, most likely to be detected by PLATO and *Kepler*, may be the inner pairs of triples and other hierarchies, and a further aim is to improve our understanding of the mass ratio distribution of binaries with $P < 10$ days. PLATO data, when available, will greatly enhance our understanding of binaries and other hierarchies, and the work presented here may serve as a guide to the benefits that are to come.

An essential part of this project has been deriving the estimated intrinsic exoplanet distributions. The two estimated distributions are likely to bracket the true distribution, and provide insights into the likely intrinsic distribution at planet radii $0.0 < log\ R/R_{\oplus} < 0.2$, where both estimated distributions converge. As highlighted in Chapter 6 and in Section 8.1.1, the two distributions have very different normalisation factors of 15 and 1.5, representing the differences in the assumed *Kepler* sensitivity on which the distributions were based. Understanding the true normalisation factor and sensitivity required to match the *Kepler* catalogue of confirmed planets will be essential in moving this part of this work forward.

Kepler, the most significant transiting exoplanet survey to date, is a rich source of data on exoplanets and on false positives. The *Kepler* data archives have been mined to calibrate

expected numbers of planets and of eclipsing binaries detectable in the forthcoming PLATO mission. The NASA Exoplanet Science archive (NExSci) holds data on confirmed planets and on false positives. This is complemented by data held in the *Kepler* Eclipsing Binary Catalogue (KEBC), which identifies more unblended eclipsing binaries than there are confirmed planets listed on NExSci.

Synthetic fields were generated using the population synthesis code, `BiSEPS`. *Kepler* data, relating both to the observable exoplanet and eclipsing binary populations, was then used to calibrate the synthetic exoplanet and eclipsing binary populations seeded in these fields. Three types of synthetic field were used:

1. Fields of 1 deg^2 centred on $l = 50^\circ$, $b = 8^\circ$ were generated with eleven different initial mass ratio distributions, in order to calibrate the mass ratio distribution of the eclipsing binary population with $P < 10$ days to that observed in the KEBC, as described in Chapter 7.2.

2. The synthetic *Kepler* field described in Farmer et al. (2013), which had been calibrated to the original *Kepler* Input Catalogue and was magnitude limited to $Kp \leq 16$, was used to generate the intrinsic exoplanet distributions, as described in Chapter 6 and to further calibrate the eclipsing binary population by period, as described in Chapter 7.

3. PLATO simulations of the two proposed Long Look fields were generated on the same basis as the synthetic *Kepler* field described in Farmer et al. (2013).

Table 4.1 summarises all fields used in this work, their purpose and the chapters in which they were used.

Synthetic populations of exoplanets and eclipsing binaries were generated, based on the calibrations described above. These synthetic populations were compared in order to estimate the ratio of exoplanets to eclipsing binaries both in the PLATO priority classes and in the general population.

The methods used in this work can be extended to other transiting exoplanet surveys, by incorporating the detection parameters of the relevant observatory into the code.

9.1.2 Main findings

The main conclusions to be drawn from the results presented in Chapter 6, Chapter 7 and Chapter 8 are:

1. For the PLATO priority populations, the Long Look South field is likely to contain more observable planets than the Long Look North field.

2. The Long Look South field appears to be less contaminated by background eclipsing binaries than the Long Look North field.

3. When considering unblended eclipsing binaries, however, the Long Look South field appears to be more contaminated than the Long Look North field.

4. Giant planets whose host is a target star are affected by dilution due to blending with other stars in the same pixel in a *Kepler*-like population ($V < 16$, F5-M), but in the PLATO P1, P2 and P5 priority classes (F5-K and $V < 8$, $V < 11$ and $V < 13$ respectively) this is not a concern. However, dilution by blending does affect smaller planets (Neptune-like and terrestrial), leading to an excess of planets with an apparent radius of $-0.2 < log\ R/R_{\oplus} < 0.0$ compared to the true number of planets of that size in all populations except P2 and P4.

5. Regions of the synthetic Long Look North field were compared by magnitude distribution to both the current *Kepler* Input Catalogue, hosted on MAST, and to Gaia DR1. The synthetic field was found to match the data on MAST, while Gaia DR1 showed only 60% of the expected number of systems, compared both to MAST and BiSEPS. Considerations of the completeness of Gaia DR1 are recognised (Gaia Collaboration et al., 2016) and are expected to be improved in forthcoming data releases, including Gaia DR2 (April 2018).

Note that no conclusions are drawn regarding dilution by blending of planets around stars which are not the targets in the observations and which, following *Kepler*, should properly be regarded as false positives as the planet host was not the star which was being observed. The magnitude limits set for the P1, P2 and P5 populations should mitigate against the risk of blended giant planets around nearby stars being mistaken for terrestrial planets orbiting the target star. As evidenced in the UKIRT images on the *Kepler* Exoplanet Follow Up pages, or the spectroscopic data in the Gaia data releases, the brighter the star, the more likely it is that nearby stars are difficult to identify. Also, as shown in Eq. 1.14, the greater the magnitude difference between the background star and the target star, the less likely it is that a background giant planet will be observable after blending.

For the Long Look fields, two observing strategies are being considered[1]: two years each for Long Look South and Long Look North, or three years for Long Look South and one year for Long Look North. These interim results suggest a strong preference for the 3+1 (LLS + LLN) strategy, especially as this is probably the only way to obtain the three detections required to confirm a true Earth twin (one in each year in LLS).

My results indicate that PLATO is more than capable of detecting an Earth twin. However, the question of whether a signal believed to be from a transiting terrestrial exoplanet is more likely to be from a rocky planet orbiting the target star or from a contaminant, such as a background eclipsing binary, is highly sensitive to the true intrinsic exoplanet distribution.

This is an interim result with considerable uncertainties, and steps that may improve our understanding of false positives and shallow eclipsing binaries in transiting exoplanet surveys are described in the following section.

[1]PLATO Science Conference, University of Warwick, September 2017, https://warwick.ac.uk/fac/sci/physics/research/astro/research/meetings/plato_mission_conference2017

9.2 Future Work

9.2.1 Development of BiSEPS stellar models

The following developments of the stellar models used in BiSEPS would strengthen the code's ability to simulate a realistic synthetic Galactic population.

At present, BiSEPS uses fitting formulae for its stellar models. The next step would be to incorporate stellar tracks from a stellar evolutionary code, such as MESA (Paxton et al. (2011), Paxton et al. (2013), Paxton et al. (2015)). This would require running MESA for both single star and binary models of given initial parameters. From the stellar evolutionary tracks, the data currently noted from the fitting formulae of Hurley et al. (2000) and Hurley et al. (2002), and required by the population synthesis part of the model, would need to be recorded. Stellar properties include the mass, radius, luminosity and effective temperature of both components in a binary system, as well as orbital period, metallicity and the age at the start and end of each model in the sequence. Also recorded are a series of keywords which indicate the evolutionary status of the two components, whether mass transfer has ever occurred, if mass transfer is currently occurring and if so, of what kind. Given the resolution that would be required to achieve a comprehensive set of stellar tracks, this would be a considerable undertaking in terms of computational time and storage space, but the benefits would be significant.

A more realistic consideration of eccentricity would also bring significant benefits. At present, the code makes the simplifying assumption that $e = 0$ at all times, which is not representative of the observed binary properties in the Galaxy. A more appropriate physical model follows the evolution of e due to tidal and orbital interactions, from its initial value to circularisation. Eccentricity has implications for eclipse duration and for the timing of the secondary eclipse in relation to the primary eclipse, and also requires an understanding of the relationship between the line of sight to the observer, periastron and apastron. With a code that is not limited to $e = 0$ it would be possible to more closely reproduce the light curves of observed eclipsing binaries, further enhancing the potential of BiSEPS as a predictive tool. Eccentricity may also impact on the duration of pre-mass transfer evolutionary phases.

Another useful enhancement would be to continue evolution of systems beyond the common envelope phase where the result is a merger. Common envelope evolution is an intrinsically 3D problem that can't be tackled with 1D stellar codes. There is a large variation in both temporal and spatial scale involved so that 3D hydro codes are also limited and cannot resolve all the phases of the common envelope. A possible solution would be to incorporate evolutionary tracks of merged systems from MESA. Pending the availability of such tracks, one other option may be to utilise results from the "Make Me A Star" code[2], [3],

[2]http://adsabs.harvard.edu/abs/2014ascl.soft12010L
[3]http://starsmasher.allegheny.edu/jalombar/mmas/

which synthesises bodies resulting from stellar collisions (Lombardi et al., 2003), although mergers and collisions are likely to have different outcomes.

9.2.2 Development of BiSEPS synthetic population

Work to date has concentrated on the two proposed Long Look fields. A key question for those working on the position of the Long Look fields is, can observations move closer to the Galactic plane? In this light, simulating a population closer to the Galactic plane would be extremely useful.

The positions of the Long Look fields are not yet fixed, and potential Step & Stare fields also need to covered by eclipsing binary and transiting exoplanet simulations. In this light, extending the simulations to other parts of the sky would be highly informative. Fig. 4.1 indicates the regions for which synthetic fields have already been simulated, covering large areas of the sky, which can already be used in synthetic exoplanet and synthetic eclipsing binary work. The biggest need, moving forward, is simulations closer to the Galactic plane, which will be expensive both in terms of computational time and disc space.

The consistency and validity of the simulations must also always be verified using Gaia DR2 (April 2018) and future releases, scheduled for mid to late 2020 (DR3) and the end of 2022 (DR4),[4] to enable increasingly accurate calibration of the BiSEPS simulations. Of particular interest will be DR3, which will include data on binaries.

Given that my simulations indicate that most hosts of planets detectable by PLATO are < 1.2 kpc from Earth, the local stellar density, which cannot be assumed to be isotropic, and local extinction are also highly relevant. As more parallaxes, and hence distance estimates, are included in the Gaia data, an enhanced understanding of the solar neighbourhood will become available. This will inform future iterations of these simulations.

The distribution of binaries is also relevant. Observations indicate that the binary fraction is larger for high-mass stars, but in my simulations binaries are distributed equally regardless of mass. The calibration by initial mass ratio may be correcting for this, as it is noticeable in my simulations that, while the uncalibrated results indicate the number of binaries begins to fall as the apparent radius approaches solar radius, the calibrated results indicate that the number of binaries continues to rise. Achieving a closer match in the initially simulated population to the observed population will increase the accuracy of the results obtainable with BiSEPS.

Neither the halo population nor the bulge is currently included in my BiSEPS simulations. While it is unlikely that stars in the bulge will contaminate PLATO observations in the Long Look fields, it is theoretically possible that contaminating eclipsing binaries detectable by

[4]https://www.cosmos.esa.int/web/gaia/release

PLATO may be contained within globular clusters in the halo. Inclusion of this option in simulations, especially where verified against observations, would be useful, although this is of lower priority than the other topics listed above.

9.2.3 Development of exoplanet models

The characteristic distributions 'zeta' and 'omega' are regarded as limiting cases for the true intrinsic exoplanet distribution. They can be improved as follows.

In the same way that DR25 revised the radii of planets on NExSci, future spectroscopy and asteroseismology will, through an improved understanding of the stellar parameters and/or an increase in the number of confirmed planets following work in the community, result in revised planetary statistics, including the distribution of planets over radius and period. Observations from space-based observatories over long time scales, such as the planned TESS Continuous Viewing Zone, will improve our understanding of the distribution of planets in the period-radius plane. Monitoring of this rapidly expanding field is therefore a necessity.

Currently we consider exoplanet frequency over a function of radius and period only. Going forward, calibrating the type of planet by the host star, for example considering the low occurrence rate of giant planets around M-class stars, would result in a more finely tuned exoplanet calibration. The intrinsic exoplanet distributions also do not consider planetary system architecture, as they use an isotropic distribution by angle of inclination. Compensating for this would allow the simulation of multiplanet systems. Resonances in short period systems could also be considered.

Incorporating eccentricity of planetary orbits into the simulations would open up the possibility of matching simulated to observed planets by transit duration, planet host and transit depth, in the same way that short period binaries were matched in Chapter 7.

A growing body of work strongly hints at dilution by blending, with bound and unbound companions, of planets in the *Kepler* field (Ciardi et al. (2015), Furlan et al. (2017)). As well as considering background contaminants, investigation of planets in both P-type and S-type orbits in binary systems, both observed and simulated, will further improve understanding of the intrinsic exoplanet distribution.

Blending of planet hosts with unbound companions in the PLATO simulations is, at present, only considered within the pixel within which the synthetic planet host is situated. Including blending with stars in neighbouring pixels will enhance the PLATO simulations and lead to a more accurate understanding of the predicted planet to binary ratio.

9.2.4 Development of eclipsing binary models

Unrecognised stellar multiplicity may arise from triple systems. Such systems are presently not included in BiSEPS, which only considers single star and binary populations. Extending the synthetic population to consider triples and other hierarchies will assist in calibrating the Galactic model.

The inclusion of eccentricity in the BiSEPS stellar models, also referred to above, is also an essential step in accurately reproducing the observed population.

The synthetic eclipsing binary population, calibrated to include those binary systems which may be observed in a transiting exoplanet survey, resulted in an excess of binaries at periods $P < 6$ days and a deficit of binaries at 10-40 days. Further work to understand this divergence would enhance the credibility of the binary simulations and would contribute to our understanding of binary evolution and formation.

9.2.5 Inclusion of other sources of false positives

While this study has focused on false positives due to eclipsing binaries, other types of false positives and false alarms can be expected to affect transiting exoplanet surveys.

Some sources of false alarms noted in the *Kepler* data, such as cosmic ray strikes, are unlikely to affect all cameras in a group onboard PLATO at the same time, so are unlikely to be as significant a problem for PLATO. Other sources, such as instrumental effects, would have been unique to *Kepler*; however, as PLATO is developed and tested, instrumental effects unique to PLATO may become apparent.

Christiansen et al. (2012) demonstrated that, by the use of different integration times, the *Kepler* pipeline could identify variable stars, and it is assumed that PLATO software similarly will be able to identify variable target stars. However, when only the wings of the PSF of variable stars and distant eclipsing binaries are seen, these can mimic planets, particularly terrestrial planets. To assess the likelihood of contamination from a distant star, consideration of separation in terms of rows and columns within a CCD is likely to be required, just as it is for *Kepler*. The PLATO Definition Study Report states that the 24 cameras are expected to be co-aligned. The Report indicates that misalignment between cameras should be minimised and that the maximum misalignment for each camera should be ≈ 3.4', which would ensure that less than 1% of the stars in the field of view would be lost. The concern expressed in the PLATO Definition Study Report is that misalignment would decrease the overlap of the different groups and of the cameras within a group, leading to loss of coverage of stars around the edges. On this basis, misalignment may also reduce the likelihood of a false positive arising from distant eclipsing binaries or variable stars. Misalignment may occur at any point following assembly of the satellite, including during launch and during flight to the L2

position, from where PLATO will survey the Galaxy. Ephemeris matches to distant stars and consideration of the effects of varying degrees of misalignment on contamination by nearby contaminants should be examined as the work begun in this project progresses.

9.3 Summary: Chapter 9

The research undertaken in this project has provided insight into the nature of eclipsing binaries that may contaminate transiting exoplanet surveys, as well as the nature of the intrinsic exoplanet distribution. Estimates of detectable planets and observable eclipsing binaries, with various degrees of blending, indicate that if the degree of blending can be limited, more terrestrial planets than binaries will be identified by PLATO. The future work described in this chapter will be invaluable in enhancing the reliability of these estimates.

Simulations cannot be conducted without reference to observations. Key to developing this work further is calibration to data provided by Gaia DR2 and later Gaia data releases. This will indicate if the increased numbers of planets (and unblended eclipsing binaries) recorded in the synthetic Southern field is a robust feature of the model.

References

Abdul-Masih, M., Prša, A., Conroy, K., Bloemen, S., Boyajian, T., Doyle, L. R., Johnston, C., Kostov, V., Latham, D. W., Matijevič, G., Shporer, A., and Southworth, J. (2016). Kepler Eclipsing Binary Stars. VIII. Identification of False Positive Eclipsing Binaries and Re-extraction of New Light Curves. *The Astronomical Journal*, 151:101.

Adams, E. R., Seager, S., and Elkins-Tanton, L. (2008). Ocean Planet or Thick Atmosphere: On the Mass-Radius Relationship for Solid Exoplanets with Massive Atmospheres. *The Astrophysical Journal*, 673:1160–1164.

Alsubai, K. A., Parley, N. R., Bramich, D. M., Horne, K., Collier Cameron, A., West, R. G., Sorensen, P. M., Pollacco, D., Smith, J. C., and Fors, O. (2013). The Qatar Exoplanet Survey. *Acta Astronomica*, 63:465–480.

Andrae, R., Fouesneau, M., Creevey, O., Ordenovic, C., Mary, N., Burlacu, A., Chaoul, L., Jean-Antoine-Piccolo, A., Kordopatis, G., Korn, A., Lebreton, Y., Panem, C., Pichon, B., Thevenin, F., Walmsley, G., and Bailer-Jones, C. A. L. (2018). Gaia Data Release 2: first stellar parameters from Apsis. *ArXiv e-prints*.

Anglada-Escudé, G., Amado, P. J., Barnes, J., Berdiñas, Z. M., Butler, R. P., Coleman, G. A. L., de La Cueva, I., Dreizler, S., Endl, M., Giesers, B., Jeffers, S. V., Jenkins, J. S., Jones, H. R. A., Kiraga, M., Kürster, M., López-González, M. J., Marvin, C. J., Morales, N., Morin, J., Nelson, R. P., Ortiz, J. L., Ofir, A., Paardekooper, S.-J., Reiners, A., Rodríguez, E., Rodríguez-López, C., Sarmiento, L. F., Strachan, J. P., Tsapras, Y., Tuomi, M., and Zechmeister, M. (2016). A terrestrial planet candidate in a temperate orbit around Proxima Centauri. *Nature*, 536:437–440.

Arenou, F., Luri, X., Babusiaux, C., Fabricius, C., Helmi, A., Robin, A. C., Vallenari, A., Blanco-Cuaresma, S., Cantat-Gaudin, T., Findeisen, K., Reylé, C., Ruiz-Dern, L., Sordo, R., Turon, C., Walton, N. A., Shih, I.-C., Antiche, E., Barache, C., Barros, M., Breddels, M., Carrasco, J. M., Costigan, G., Diakité, S., Eyer, L., Figueras, F., Galluccio, L., Heu, J., Jordi, C., Krone-Martins, A., Lallement, R., Lambert, S., Leclerc, N., Marrese, P. M., Moitinho, A., Mor, R., Romero-Gómez, M., Sartoretti, P., Soria, S., Soubiran, C., Souchay, J., Veljanoski, J., Ziaeepour, H., Giuffrida, G., Pancino, E., and Bragaglia, A. (2017). Gaia Data Release 1. Catalogue validation. *Astronomy and Astrophysics*, 599:A50.

Astudillo-Defru, N., Bonfils, X., Delfosse, X., Ségransan, D., Forveille, T., Bouchy, F., Gillon, M., Lovis, C., Mayor, M., Neves, V., Pepe, F., Perrier, C., Queloz, D., Rojo, P., Santos, N. C., and Udry, S. (2015). The HARPS search for southern extra-solar planets. XXXVI. Planetary systems and stellar activity of the M dwarfs GJ 3293, GJ 3341, and GJ 3543. *Astronomy & Astrophysics*, 575:A119.

Bakos, G., Noyes, R. W., Kovács, G., Stanek, K. Z., Sasselov, D. D., and Domsa, I. (2004). Wide-Field Millimagnitude Photometry with the HAT: A Tool for Extrasolar Planet Detection. *The Publications of the Astronomical Society of the Pacific*, 116:266–277.

Bakos, G. Á., Csubry, Z., Penev, K., Bayliss, D., Jordán, A., Afonso, C., Hartman, J. D., Henning, T., Kovács, G., Noyes, R. W., Béky, B., Suc, V., Csák, B., Rabus, M., Lázár, J., Papp, I., Sári, P., Conroy, P., Zhou, G., Sackett, P. D., Schmidt, B., Mancini, L., Sasselov, D. D., and Ueltzhoeffer, K. (2013). HATSouth: A Global Network of Fully Automated Identical Wide-Field Telescopes. *Publications of the Astronomical Society of the Pacific*, 125:154.

Ballard, S. and Johnson, J. A. (2016). The Kepler Dichotomy among the M Dwarfs: Half of Systems Contain Five or More Coplanar Planets. *The Astrophysical Journal*, 816:66.

Barban, C., Goupil, M. J., Van't Veer-Menneret, C., Garrido, R., Kupka, F., and Heiter, U. (2003). New grids of ATLAS9 atmospheres. II. Limb-darkening coefficients for the Strömgren photometric system for A-F stars. *Astronomy and Astrophysics*, 405:1095–1105.

Barclay, T., Quintana, E. V., Adams, F. C., Ciardi, D. R., Huber, D., Foreman-Mackey, D., Montet, B. T., and Caldwell, D. (2015). The Five Planets in the Kepler-296 Binary System All Orbit the Primary: A Statistical and Analytical Analysis. *The Astrophysical Journal*, 809:7.

Barnes, J. W., Linscott, E., and Shporer, A. (2011). Measurement of the Spin-Orbit Misalignment of KOI-13.01 from Its Gravity-darkened Kepler Transit Lightcurve. *The Astrophysical Journal Supplement*, 197:10.

Batalha, N. M., Borucki, W. J., Bryson, S. T., Buchhave, L. A., Caldwell, D. A., Christensen-Dalsgaard, J., Ciardi, D., Dunham, E. W., Fressin, F., Gautier, III, T. N., Gilliland, R. L., Haas, M. R., Howell, S. B., Jenkins, J. M., Kjeldsen, H., Koch, D. G., Latham, D. W., Lissauer, J. J., Marcy, G. W., Rowe, J. F., Sasselov, D. D., Seager, S., Steffen, J. H., Torres, G., Basri, G. S., Brown, T. M., Charbonneau, D., Christiansen, J., Clarke, B., Cochran, W. D., Dupree, A., Fabrycky, D. C., Fischer, D., Ford, E. B., Fortney, J., Girouard, F. R., Holman, M. J., Johnson, J., Isaacson, H., Klaus, T. C., Machalek, P., Moorehead, A. V., Morehead, R. C., Ragozzine, D., Tenenbaum, P., Twicken, J., Quinn, S., VanCleve, J., Walkowicz, L. M., Welsh, W. F., Devore, E., and Gould, A. (2011). Kepler's First Rocky Planet: Kepler-10b. *The Astrophysical Journal*, 729:27.

Batalha, N. M., Borucki, W. J., Koch, D. G., Bryson, S. T., Haas, M. R., Brown, T. M., Caldwell, D. A., Hall, J. R., Gilliland, R. L., Latham, D. W., Meibom, S., and Monet, D. G. (2010a). Selection, Prioritization, and Characteristics of Kepler Target Stars. *The Astrophysical Journal Letters*, 713:L109–L114.

Batalha, N. M., Rowe, J. F., Gilliland, R. L., Jenkins, J. J., Caldwell, D., Borucki, W. J., Koch, D. G., Lissauer, J. J., Dunham, E. W., Gautier, T. N., Howell, S. B., Latham, D. W., Marcy, G. W., and Prsa, A. (2010b). Pre-spectroscopic False-positive Elimination of Kepler Planet Candidates. *The Astrophysical Journal Letters*, 713:L103–L108.

Beaugé, C. and Nesvorný, D. (2012). Multiple-planet Scattering and the Origin of Hot Jupiters. *The Astrophysical Journal*, 751:119.

Beaulieu, J.-P., Bennett, D. P., Fouqué, P., Williams, A., Dominik, M., Jørgensen, U. G., Kubas, D., Cassan, A., Coutures, C., Greenhill, J., Hill, K., Menzies, J., Sackett, P. D., Albrow, M., Brillant, S., Caldwell, J. A. R., Calitz, J. J., Cook, K. H., Corrales, E., Desort, M., Dieters, S., Dominis, D., Donatowicz, J., Hoffman, M., Kane, S., Marquette, J.-B., Martin, R., Meintjes, P., Pollard, K., Sahu, K., Vinter, C., Wambsganss, J., Woller, K., Horne, K., Steele, I., Bramich, D. M., Burgdorf, M., Snodgrass, C., Bode, M., Udalski, A., Szymański, M. K., Kubiak, M., Więckowski, T., Pietrzyński, G., Soszyński, I., Szewczyk, O., Wyrzykowski, Ł., Paczyński, B., Abe, F., Bond, I. A., Britton, T. R., Gilmore, A. C., Hearnshaw, J. B., Itow, Y., Kamiya, K., Kilmartin, P. M., Korpela, A. V., Masuda, K., Matsubara, Y., Motomura, M., Muraki, Y., Nakamura, S., Okada, C., Ohnishi, K., Rattenbury, N. J., Sako, T., Sato, S., Sasaki, M., Sekiguchi, T., Sullivan, D. J., Tristram, P. J., Yock, P. C. M., and Yoshioka, T. (2006). Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. *Nature*, 439:437–440.

Beck, P. G., Hambleton, K., Vos, J., Kallinger, T., Bloemen, S., Tkachenko, A., García, R. A., Østensen, R. H., Aerts, C., Kurtz, D. W., De Ridder, J., Hekker, S., Pavlovski, K., Mathur, S., De Smedt, K., Derekas, A., Corsaro, E., Mosser, B., Van Winckel, H., Huber, D., Degroote, P., Davies, G. R., Prša, A., Debosscher, J., Elsworth, Y., Nemeth, P., Siess, L., Schmid, V. S., Pápics, P. I., de Vries, B. L., van Marle, A. J., Marcos-Arenal, P., and Lobel, A. (2014). Pulsating red giant stars in eccentric binary systems discovered from Kepler space-based photometry. A sample study and the analysis of KIC 5006817. *Astronomy and Astrophysics*, 564:A36.

Becker, J. C., Vanderburg, A., Adams, F. C., Rappaport, S. A., and Schwengeler, H. M. (2015). WASP-47: A Hot Jupiter System with Two Additional Planets Discovered by K2. *The Astrophysical Journal Letters*, 812:L18.

Blank, D. L., Feliz, D., Collins, K. A., White, G. L., Stassun, K. G., Curtis, I. A., Hart, R., Kielkopf, J. F., Nelson, P., Relles, H., Stockdale, C., Jayawardene, B., Pennypacker, C. R., Shankland, P., Reichart, D. E., Haislip, J. B., and Kouprianov, V. V. (2018). A Multi-year Search for Transits of Proxima Centauri. I. Light Curves Corresponding to Published Ephemerides. *The Astronomical Journal*, 155:228.

Bochinski, J., Haswell, C. A., Fossati, L., Dhillon, V. S., Marsh, T. R., Nickson, E., and Wheatley, P. J. (2013). High quality fast photometry of a highly irradiated transiting exoplanet WASP-12 b. *European Planetary Science Congress 2013, held 8-13 September in London, UK. Online at: http://meetings.copernicus.org/epsc2013, id.EPSC2013-857*, 8:EPSC2013–857.

Bordé, P., Rouan, D., and Léger, A. (2003). Exoplanet detection capability of the COROT space mission. *Astronomy and Astrophysics*, 405:1137–1144.

Borucki, W. J., Koch, D., Basri, G., Batalha, N., Brown, T., Caldwell, D., Caldwell, J., Christensen-Dalsgaard, J., Cochran, W. D., DeVore, E., Dunham, E. W., Dupree, A. K., Gautier, T. N., Geary, J. C., Gilliland, R., Gould, A., Howell, S. B., Jenkins, J. M., Kondo, Y., Latham, D. W., Marcy, G. W., Meibom, S., Kjeldsen, H., Lissauer, J. J., Monet, D. G., Morrison, D., Sasselov, D., Tarter, J., Boss, A., Brownlee, D., Owen, T., Buzasi, D., Charbonneau, D., Doyle, L., Fortney, J., Ford, E. B., Holman, M. J., Seager, S., Steffen, J. H., Welsh, W. F., Rowe, J., Anderson, H., Buchhave, L., Ciardi, D., Walkowicz, L., Sherry, W., Horch, E., Isaacson, H., Everett, M. E., Fischer, D., Torres, G., Johnson, J. A., Endl, M., MacQueen, P., Bryson, S. T., Dotson, J., Haas, M., Kolodziejczak, J., Van Cleve,

J., Chandrasekaran, H., Twicken, J. D., Quintana, E. V., Clarke, B. D., Allen, C., Li, J., Wu, H., Tenenbaum, P., Verner, E., Bruhweiler, F., Barnes, J., and Prsa, A. (2010). Kepler Planet-Detection Mission: Introduction and First Results. *Science*, 327:977.

Boyd, D. R. S., de Miguel, E., Patterson, J., Wood, M. A., Barrett, D., Boardman, J., Brettman, O., Cejudo, D., Collins, D., Cook, L. M., Cook, M. J., Foote, J. L., Fried, R., Gomez, T. L., Hambsch, F.-J., Jones, J. L., Kemp, J., Koff, R., Koppelman, M., Krajci, T., Lemay, D., Martin, B., McClusky, J. V., Menzies, K., Messier, D., Roberts, G., Robertson, J., Rock, J., Sabo, R., Skillman, D., Ulowetz, J., and Vanmunster, T. (2017). A 16-yr photometric campaign on the eclipsing novalike variable DW Ursae Majoris. *Monthly Notices of the Royal Astronomical Society*, 466:3417–3433.

Broeg, C., Fortier, A., Ehrenreich, D., Alibert, Y., Baumjohann, W., Benz, W., Deleuil, M., Gillon, M., Ivanov, A., Liseau, R., Meyer, M., Oloffson, G., Pagano, I., Piotto, G., Pollacco, D., Queloz, D., Ragazzoni, R., Renotte, E., Steller, M., and Thomas, N. (2013). CHEOPS: A transit photometry mission for ESA's small mission programme. In *European Physical Journal Web of Conferences*, volume 47 of *European Physical Journal Web of Conferences*, page 03005.

Brown, D. J. A. (2014). Discrepancies between isochrone fitting and gyrochronology for exoplanet host stars? *Monthly Notices of the Royal Astronomical Society*, 442:1844–1862.

Brown, D. J. A., Collier Cameron, A., Hall, C., Hebb, L., and Smalley, B. (2011a). Are falling planets spinning up their host stars? *Monthly Notices of the Royal Astronomical Society*, 415:605–618.

Brown, T. M., Latham, D. W., Everett, M. E., and Esquerdo, G. A. (2011b). Kepler Input Catalog: Photometric Calibration and Stellar Classification. *The Astronomical Journal*, 142:112.

Bryson, S., Coughlin, J. L., Thompson, S. E., Mullally, F., Haas, M. R., Batalha, N., and Kepler Team (2017). Determining Statistically Optimal Metric Thresholds for the Final Kepler Planet Candidate Catalog. In *American Astronomical Society Meeting Abstracts*, volume 230 of *American Astronomical Society Meeting Abstracts*, page 102.03.

Buchhave, L. A., Dressing, C. D., Dumusque, X., Rice, K., Vanderburg, A., Mortier, A., Lopez-Morales, M., Lopez, E., Lundkvist, M. S., Kjeldsen, H., Affer, L., Bonomo, A. S., Charbonneau, D., Collier Cameron, A., Cosentino, R., Figueira, P., Fiorenzano, A. F. M., Harutyunyan, A., Haywood, R. D., Johnson, J. A., Latham, D. W., Lovis, C., Malavolta, L., Mayor, M., Micela, G., Molinari, E., Motalebi, F., Nascimbeni, V., Pepe, F., Phillips, D. F., Piotto, G., Pollacco, D., Queloz, D., Sasselov, D., Ségransan, D., Sozzetti, A., Udry, S., and Watson, C. (2016). A 1.9 Earth Radius Rocky Planet and the Discovery of a Non-transiting Planet in the Kepler-20 System. *The Astronomical Journal*, 152:160.

Burke, C. J., Bryson, S. T., Mullally, F., Rowe, J. F., Christiansen, J. L., Thompson, S. E., Coughlin, J. L., Haas, M. R., Batalha, N. M., Caldwell, D. A., Jenkins, J. M., Still, M., Barclay, T., Borucki, W. J., Chaplin, W. J., Ciardi, D. R., Clarke, B. D., Cochran, W. D., Demory, B.-O., Esquerdo, G. A., Gautier, III, T. N., Gilliland, R. L., Girouard, F. R., Havel, M., Henze, C. E., Howell, S. B., Huber, D., Latham, D. W., Li, J., Morehead, R. C., Morton, T. D., Pepper, J., Quintana, E., Ragozzine, D., Seader, S. E., Shah, Y., Shporer, A., Tenenbaum, P., Twicken, J. D., and Wolfgang, A. (2014). Planetary Candidates Observed by

Kepler IV: Planet Sample from Q1-Q8 (22 Months). *The Astrophysical Journal Supplement*, 210:19.

Burrows, A., Sudarsky, D., and Hubeny, I. (2006). Theory for the Secondary Eclipse Fluxes, Spectra, Atmospheres, and Light Curves of Transiting Extrasolar Giant Planets. *The Astrophysical Journal*, 650:1140–1149.

Camargo, D., Bonatto, C., and Bica, E. (2015). Tracing the Galactic spiral structure with embedded clusters. *Monthly Notices of the Royal Astronomical Society*, 450:4150–4160.

Carey, S. J., Surace, J. A., Glaccum, W. J., Ingalls, J., Krick, J., Lacy, M., Lowrance, P., Laine, S., O'Linger, J., Stauffer, J. R., Willner, S. P., Hora, J. L., Hoffmann, W. F., Ashby, M. L. N., Huang, J., Marengo, M., Pahre, M. A., Wang, Z., and Fazio, G. G. (2010). Properties of Warm IRAC Data. In *American Astronomical Society Meeting Abstracts #215*, volume 42 of *Bulletin of the American Astronomical Society*, page 571.

Chabrier, G. (2001). The Galactic Disk Mass Budget. I. Stellar Mass Function and Density. *The Astrophysical Journal*, 554:1274–1281.

Chanamé, J. and Ramírez, I. (2012). Toward Precise Ages for Single Stars in the Field. Gyrochronology Constraints at Several Gyr Using Wide Binaries. I. Ages for Initial Sample. *The Astrophysical Journal*, 746:102.

Charbonneau, D., Berta, Z. K., Irwin, J., Burke, C. J., Nutzman, P., Buchhave, L. A., Lovis, C., Bonfils, X., Latham, D. W., Udry, S., Murray-Clay, R. A., Holman, M. J., Falco, E. E., Winn, J. N., Queloz, D., Pepe, F., Mayor, M., Delfosse, X., and Forveille, T. (2009). A super-Earth transiting a nearby low-mass star. *Nature*, 462:891–894.

Christiansen, J. L., Ballard, S., Charbonneau, D., Madhusudhan, N., Seager, S., Holman, M. J., Wellnitz, D. D., Deming, D., A'Hearn, M. F., and EPOXI Team (2010). Studying the Atmosphere of the Exoplanet HAT-P-7b Via Secondary Eclipse Measurements with EPOXI, Spitzer, and Kepler. *The Astrophysical Journal,*, 710:97–104.

Christiansen, J. L., Clarke, B. D., Burke, C. J., Seader, S., Jenkins, J. M., Twicken, J. D., Catanzarite, J. D., Smith, J. C., Batalha, N. M., Haas, M. R., Thompson, S. E., Campbell, J. R., Sabale, A., and Kamal Uddin, A. (2015). Measuring Transit Signal Recovery in the Kepler Pipeline II: Detection Efficiency as Calculated in One Year of Data. *The Astrophysical Journal*, 810:95.

Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., Burke, C. J., Tenenbaum, P., Seader, S., Thompson, S. E., Barclay, T. S., Clarke, B. D., Li, J., Smith, J. C., Stumpe, M. C., Twicken, J. D., and Van Cleve, J. (2012). The Derivation, Properties, and Value of Kepler's Combined Differential Photometric Precision. *Publications of the Astronomical Society of the Pacific*, 124:1279.

Ciardi, D. R., Beichman, C. A., Horch, E. P., and Howell, S. B. (2015). Understanding the Effects of Stellar Multiplicity on the Derived Planet Radii from Transit Surveys: Implications for Kepler, K2, and TESS. *The Astrophysical Journal*, 805:16.

Claret, A. (1998). Comprehensive tables for the interpretation and modeling of the light curves of eclipsing binaries. *Astronomy and Astrophysics Supplement*, 131:395–400.

Claret, A. (2000). A new non-linear limb-darkening law for LTE stellar atmosphere models. Calculations for -5.0 <= log[M/H] <= +1, 2000 K <= T_{eff} <= 50000 K at several surface gravities. *Astronomy and Astrophysics*, 363:1081–1190.

Claret, A. and Bloemen, S. (2011). Gravity and limb-darkening coefficients for the Kepler, CoRoT, Spitzer, uvby, UBVRIJHK, and Sloan photometric systems. *Astronomy and Astrophysics*, 529:A75.

Clarkson, W., Sahu, K., Anderson, J., Smith, T. E., Brown, T. M., Rich, R. M., Casertano, S., Bond, H. E., Livio, M., Minniti, D., Panagia, N., Renzini, A., Valenti, J., and Zoccali, M. (2008). Stellar Proper Motions in the Galactic Bulge from Deep Hubble Space Telescope ACS WFC Photometry. *The Astrophysical Journal*, 684:1110–1142.

Conroy, K. E., Prša, A., Stassun, K. G., Bloemen, S., Parvizi, M., Quarles, B., Boyajian, T., Barclay, T., Shporer, A., Latham, D. W., and Abdul-Masih, M. (2014a). Kepler Eclipsing Binary Stars. V. Identification of 31 Candidate Eclipsing Binaries in the K2 Engineering Dataset. *Publications of the Astronomical Society of Pacific*, 126:914.

Conroy, K. E., Prša, A., Stassun, K. G., Orosz, J. A., Fabrycky, D. C., and Welsh, W. F. (2014b). Kepler Eclipsing Binary Stars. IV. Precise Eclipse Times for Close Binaries and Identification of Candidate Three-body Systems. *The Astronomical Journal*, 147:45.

Coughlin, J. L. (2017). Planet Detection Metrics: Robovetter Completeness and Effectiveness for Data Release 25. Technical report, NASA.

Coughlin, J. L., Mullally, F., Thompson, S. E., Rowe, J. F., Burke, C. J., Latham, D. W., Batalha, N. M., Ofir, A., Quarles, B. L., Henze, C. E., Wolfgang, A., Caldwell, D. A., Bryson, S. T., Shporer, A., Catanzarite, J., Akeson, R., Barclay, T., Borucki, W. J., Boyajian, T. S., Campbell, J. R., Christiansen, J. L., Girouard, F. R., Haas, M. R., Howell, S. B., Huber, D., Jenkins, J. M., Li, J., Patil-Sabale, A., Quintana, E. V., Ramirez, S., Seader, S., Smith, J. C., Tenenbaum, P., Twicken, J. D., and Zamudio, K. A. (2016). Planetary Candidates Observed by Kepler. VII. The First Fully Uniform Catalog Based on the Entire 48-month Data Set (Q1-Q17 DR24). *The Astrophysical Journal Supplement Series*, 224:12.

Coughlin, J. L., Thompson, S. E., Bryson, S. T., Burke, C. J., Caldwell, D. A., Christiansen, J. L., Haas, M. R., Howell, S. B., Jenkins, J. M., Kolodziejczak, J. J., Mullally, F. R., and Rowe, J. F. (2014). Contamination in the Kepler Field. Identification of 685 KOIs as False Positives via Ephemeris Matching Based on Q1-Q12 Data. *The Astronomical Journal*, 147:119.

Davis, P. J., Kolb, U., and Willems, B. (2010). A comprehensive population synthesis study of post-common envelope binaries. *Monthly Notices of the Royal Astronomical Society*, 403:179–195.

Davis, P. J., Kolb, U., Willems, B., and Gänsicke, B. T. (2008). How many cataclysmic variables are crossing the period gap? A test for the disruption of magnetic braking. *Monthly Notices of the Royal Astronomical Society*, 389:1563–1576.

Deming, D., Harrington, J., Seager, S., and Richardson, L. J. (2006). Strong Infrared Emission from the Extrasolar Planet HD 189733b. *The Astrophysical Journal*, 644:560–564.

Demory, B.-O., Gillon, M., Deming, D., Valencia, D., Seager, S., Benneke, B., Lovis, C., Cubillos, P., Harrington, J., Stevenson, K. B., Mayor, M., Pepe, F., Queloz, D., Ségransan, D., and Udry, S. (2011). Detection of a transit of the super-Earth 55 Cancri e with warm Spitzer. *Astronomy & Astrophysics*, 533:A114.

Dotter, A., Chaboyer, B., Jevremović, D., Kostov, V., Baron, E., and Ferguson, J. W. (2008). The Dartmouth Stellar Evolution Database. *The Astrophysical Journal Supplement Series*, 178:89–101.

Drimmel, R., Cabrera-Lavers, A., and López-Corredoira, M. (2003). A three-dimensional Galactic extinction model. *Astronomy and Astrophysics*, 409:205–215.

Duquennoy, A. and Mayor, M. (1991). Multiplicity among solar-type stars in the solar neighbourhood. II - Distribution of the orbital elements in an unbiased sample. *Astronomy and Astrophysics*, 248:485–524.

Farmer, R., Kolb, U., and Norton, A. J. (2013). The true stellar parameters of the Kepler target list. *Monthly Notices of the Royal Astronomical Society*, 433:1133–1145.

Fischer, D. A. and Marcy, G. W. (1992). Multiplicity among M dwarfs. *Astrophysical Journal*, 396:178–194.

Fressin, F., Torres, G., Charbonneau, D., Bryson, S. T., Christiansen, J., Dressing, C. D., Jenkins, J. M., Walkowicz, L. M., and Batalha, N. M. (2013). The False Positive Rate of Kepler and the Occurrence of Planets. *The Astrophysical Journal*, 766:81.

Fuhrmann, K. (1998). Nearby stars of the Galactic disk and halo. *Astronomy and Astrophysics*, 338:161–183.

Fuller, J. (2017). Heartbeat stars, tidally excited oscillations and resonance locking. *Monthly Notices of the Royal Astronomical Society*, 472:1538–1564.

Fulton, B. J., Petigura, E. A., Howard, A. W., Isaacson, H., Marcy, G. W., Cargile, P. A., Hebb, L., Weiss, L. M., Johnson, J. A., Morton, T. D., Sinukoff, E., Crossfield, I. J. M., and Hirsch, L. A. (2017). The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets. *ArXiv e-prints*.

Furlan, E., Ciardi, D. R., Everett, M. E., Saylors, M., Teske, J. K., Horch, E. P., Howell, S. B., van Belle, G. T., Hirsch, L. A., Gautier, III, T. N., Adams, E. R., Barrado, D., Cartier, K. M. S., Dressing, C. D., Dupree, A. K., Gilliland, R. L., Lillo-Box, J., Lucas, P. W., and Wang, J. (2017). The Kepler Follow-up Observation Program. I. A Catalog of Companions to Kepler Stars from High-Resolution Imaging. *The Astronomical Journal*, 153:71.

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Mignard, F., Drimmel, R., Babusiaux, C., Bailer-Jones, C. A. L., Bastian, U., and et al. (2016). Gaia Data Release 1. Summary of the astrometric, photometric, and survey properties. *Astronomy & Astrophysics*, 595:A2.

Gaidos, E., Mann, A. W., Kraus, A. L., and Ireland, M. (2016). They are small worlds after all: revised properties of Kepler M dwarf stars and their planets. *Monthly Notices of the Royal Astronomical Society*, 457:2877–2899.

García, R. A., Ceillier, T., Salabert, D., Mathur, S., van Saders, J. L., Pinsonneault, M., Ballot, J., Beck, P. G., Bloemen, S., Campante, T. L., Davies, G. R., do Nascimento, Jr., J.-D., Mathis, S., Metcalfe, T. S., Nielsen, M. B., Suárez, J. C., Chaplin, W. J., Jiménez, A., and Karoff, C. (2014). Rotation and magnetism of Kepler pulsating solar-like stars. Towards asteroseismically calibrated age-rotation relations. *Astronomy and Astrophysics*, 572:A34.

Getley, A. K., Carter, B., King, R., and O'Toole, S. (2017). Evidence for a planetary mass third body orbiting the binary star KIC 5095269. *Monthly Notices of the Royal Astronomical Society*, 468:2932–2937.

Gilliland, R. L., Chaplin, W. J., Jenkins, J. M., Ramsey, L. W., and Smith, J. C. (2015). Kepler Mission Stellar and Instrument Noise Properties Revisited. *The Astronomical Journal*, 150:133.

Gilliland, R. L., Marcy, G. W., Rowe, J. F., Rogers, L., Torres, G., Fressin, F., Lopez, E. D., Buchhave, L. A., Christensen-Dalsgaard, J., Désert, J.-M., Henze, C. E., Isaacson, H., Jenkins, J. M., Lissauer, J. J., Chaplin, W. J., Basu, S., Metcalfe, T. S., Elsworth, Y., Handberg, R., Hekker, S., Huber, D., Karoff, C., Kjeldsen, H., Lund, M. N., Lundkvist, M., Miglio, A., Charbonneau, D., Ford, E. B., Fortney, J. J., Haas, M. R., Howard, A. W., Howell, S. B., Ragozzine, D., and Thompson, S. E. (2013). Kepler-68: Three Planets, One with a Density between that of Earth and Ice Giants. *The Astrophysical Journal,*, 766:40.

Gillon, M., Triaud, A. H. M. J., Demory, B.-O., Jehin, E., Agol, E., Deck, K. M., Lederer, S. M., de Wit, J., Burdanov, A., Ingalls, J. G., Bolmont, E., Leconte, J., Raymond, S. N., Selsis, F., Turbet, M., Barkaoui, K., Burgasser, A., Burleigh, M. R., Carey, S. J., Chaushev, A., Copperwheat, C. M., Delrez, L., Fernandes, C. S., Holdsworth, D. L., Kotze, E. J., Van Grootel, V., Almleaky, Y., Benkhaldoun, Z., Magain, P., and Queloz, D. (2017). Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1. *Nature*, 542:456–460.

Gilmore, G., Wyse, R. F. G., and Jones, J. B. (1995). A determination of the thick disk chemical abundance distribution: Implications for galaxy evolution. *Astronomical Journal*, 109:1095–1111.

Girardi, L., Barbieri, M., Groenewegen, M. A. T., Marigo, P., Bressan, A., Rocha-Pinto, H. J., Santiago, B. X., Camargo, J. I. B., and da Costa, L. N. (2012). TRILEGAL, a TRIdimensional modeL of thE GALaxy: Status and Future. *Astrophysics and Space Science Proceedings*, 26:165.

Girardi, L., Bertelli, G., Bressan, A., Chiosi, C., Groenewegen, M. A. T., Marigo, P., Salasnich, B., and Weiss, A. (2002). Theoretical isochrones in several photometric systems. I. Johnson-Cousins-Glass, HST/WFPC2, HST/NICMOS, Washington, and ESO Imaging Survey filter sets. *Astronomy and Astrophysics*, 391:195–212.

Girardi, L., Bressan, A., Bertelli, G., and Chiosi, C. (2000). Evolutionary tracks and isochrones for low- and intermediate-mass stars: From 0.15 to 7 M_{sun}, and from Z=0.0004 to 0.03. *Astronomy and Astrophysics Supplement*, 141:371–383.

Girardi, L., Grebel, E. K., Odenkirchen, M., and Chiosi, C. (2004). Theoretical isochrones in several photometric systems. II. The Sloan Digital Sky Survey ugriz system. *Astronomy and Astrophysics*, 422:205–215.

Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., and da Costa, L. (2005). Star counts in the Galaxy. Simulating from very deep to very shallow photometric surveys with the TRILEGAL code. *Astronomy and Astrophysics*, 436:895–915.

Gratier, P., Bron, E., Gerin, M., Pety, J., Guzman, V. V., Orkisz, J., Bardeau, S., Goicoechea, J. R., Le Petit, F., Liszt, H., Öberg, K., Peretto, N., Roueff, E., Sievers, A., and Tremblin, P. (2017). Dissecting the molecular structure of the Orion B cloud: insight from principal component analysis. *Astronomy & Astrophysics*, 599:A100.

Hambleton, K., Kurtz, D. W., Prša, A., Quinn, S. N., Fuller, J., Murphy, S. J., Thompson, S. E., Latham, D. W., and Shporer, A. (2016). KIC 3749404: a heartbeat star with rapid apsidal advance indicative of a tertiary component. *Monthly Notices of the Royal Astronomical Society*, 463:1199–1212.

Haswell, C. A. (2010). *Transiting Exoplanets*. Cambridge University Press.

Haswell, C. A., Fossati, L., Ayres, T., France, K., Froning, C. S., Holmes, S., Kolb, U. C., Busuttil, R., Street, R. A., Hebb, L., Collier Cameron, A., Enoch, B., Burwitz, V., Rodriguez, J., West, R. G., Pollacco, D., Wheatley, P. J., and Carter, A. (2012). Near-ultraviolet Absorption, Chromospheric Activity, and Star-Planet Interactions in the WASP-12 system. *The Astrophysical Journal*, 760:79.

Hayashi, C., Nakazawa, K., and Adachi, I. (1977). Long-Term Behavior of Planetesimals and the Formation of the Planets. *Publications of the Astronomical Society of Japan*, 29:163–196.

Haywood, M. (2001). A revision of the solar neighbourhood metallicity distribution. *Monthly Notices of the Royal Astronomical Society*, 325:1365–1382.

Hébrard, G., Lecavelier Des Étangs, A., Vidal-Madjar, A., Désert, J.-M., and Ferlet, R. (2004). Evaporation Rate of Hot Jupiters and Formation of Chthonian Planets. In Beaulieu, J., Lecavelier Des Etangs, A., and Terquem, C., editors, *Extrasolar Planets: Today and Tomorrow*, volume 321 of *Astronomical Society of the Pacific Conference Series*, page 203.

Hełminiak, K. G., Ukita, N., Kambe, E., Kozłowski, S. K., Pawłaszek, R., Maehara, H., Baranec, C., and Konacki, M. (2017). KIC 4150611: a rare multi-eclipsing quintuple with a hybrid pulsator. *ArXiv e-prints*.

Henry, G. W., Marcy, G. W., Butler, R. P., and Vogt, S. S. (2000). A Transiting "51 Peg-like" Planet. *The Astrophysical Journal*, 529:L41–L44.

Hernán-Obispo, M., Tuomi, M., Gálvez-Ortiz, M. C., Golovin, A., Barnes, J. R., Jones, H. R. A., Kane, S. R., Pinfield, D., Jenkins, J. S., Petit, P., Anglada-Escudé, G., Marsden, S. C., Catalán, S., Jeffers, S. V., de Castro, E., Cornide, M., Garcés, A., Jones, M. I., Gorlova, N., and Andreev, M. (2015). Analysis of combined radial velocities and activity of BD+20 1790: evidence supporting the existence of a planetary companion. *Astronomy and Astrophysics*, 576:A66.

Hirsch, L. A., Ciardi, D. R., Howard, A. W., Everett, M. E., Furlan, E., Saylors, M., Horch, E. P., Howell, S. B., Teske, J., and Marcy, G. W. (2017). Assessing the Effect of Stellar Companions from High-resolution Imaging of Kepler Objects of Interest. *The Astronomical Journal*, 153:117.

Hong, K., Lee, J. W., Lee, D.-J., Kim, S.-L., Koo, J.-R., Park, J.-H., Lee, C.-U., Kim, D.-J., Cha, S.-M., and Lee, Y. (2017). Photometric Properties of the HW Vir-type Binary OGLE-GD-ECL-11388. *Publications of the Astronomical Society of the Pacific*, 129(1):014202.

Howard, A. W., Marcy, G. W., Bryson, S. T., Jenkins, J. M., Rowe, J. F., Batalha, N. M., Borucki, W. J., Koch, D. G., Dunham, E. W., Gautier, III, T. N., Van Cleve, J., Cochran, W. D., Latham, D. W., Lissauer, J. J., Torres, G., Brown, T. M., Gilliland, R. L., Buchhave, L. A., Caldwell, D. A., Christensen-Dalsgaard, J., Ciardi, D., Fressin, F., Haas, M. R., Howell, S. B., Kjeldsen, H., Seager, S., Rogers, L., Sasselov, D. D., Steffen, J. H., Basri, G. S., Charbonneau, D., Christiansen, J., Clarke, B., Dupree, A., Fabrycky, D. C., Fischer, D. A., Ford, E. B., Fortney, J. J., Tarter, J., Girouard, F. R., Holman, M. J., Johnson, J. A., Klaus, T. C., Machalek, P., Moorhead, A. V., Morehead, R. C., Ragozzine, D., Tenenbaum, P., Twicken, J. D., Quinn, S. N., Isaacson, H., Shporer, A., Lucas, P. W., Walkowicz, L. M., Welsh, W. F., Boss, A., Devore, E., Gould, A., Smith, J. C., Morris, R. L., Prsa, A., Morton, T. D., Still, M., Thompson, S. E., Mullally, F., Endl, M., and MacQueen, P. J. (2012). Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler. *The Astrophysical Journal Supplement*, 201:15.

Howell, S. B., Sobeck, C., Haas, M., Still, M., Barclay, T., Mullally, F., Troeltzsch, J., Aigrain, S., Bryson, S. T., Caldwell, D., Chaplin, W. J., Cochran, W. D., Huber, D., Marcy, G. W., Miglio, A., Najita, J. R., Smith, M., Twicken, J. D., and Fortney, J. J. (2014). The K2 Mission: Characterization and Early Results. *Publications of the Astronomical Society of the Pacific*, 126:398.

Hsu, D. C., Ford, E. B., Ragozzine, D., and Morehead, R. C. (2018). Improving the Accuracy of Planet Occurrence Rates from Kepler using Approximate Bayesian Computation. *ArXiv e-prints*.

Hurley, J. R., Pols, O. R., and Tout, C. A. (2000). Comprehensive analytic formulae for stellar evolution as a function of mass and metallicity. *Monthly Notices of the Royal Astronomical Society*, 315:543–569.

Hurley, J. R., Tout, C. A., and Pols, O. R. (2002). Evolution of binary stars and the effect of tides on binary populations. *Monthly Notices of the Royal Astronomical Society*, 329:897–928.

Isella, A., Guidi, G., Testi, L., Liu, S., Li, H., Li, S., Weaver, E., Boehler, Y., Carperter, J. M., De Gregorio-Monsalvo, I., Manara, C. F., Natta, A., Pérez, L. M., Ricci, L., Sargent, A., Tazzari, M., and Turner, N. (2016). Ringed Structures of the HD 163296 Protoplanetary Disk Revealed by ALMA. *Physical Review Letters*, 117(25):251101.

Jackson, B., Barnes, R., and Greenberg, R. (2009). Observational Evidence for Tidal Destruction of Exoplanets. *The Astrophysical Journal*, 698:1357–1366.

Jeans, J. H. (1902). The Stability of a Spherical Nebula. *Philosophical Transactions of the Royal Society of London Series A*, 199:1–53.

Jehin, E., Gillon, M., Queloz, D., Magain, P., Manfroid, J., Chantry, V., Lendl, M., Hutsemékers, D., and Udry, S. (2011). TRAPPIST: TRAnsiting Planets and PlanetesImals Small Telescope. *The Messenger*, 145:2–6.

Jenkins, J. S., Tuomi, M., Brasser, R., Ivanyuk, O., and Murgas, F. (2013). Two Super-Earths Orbiting the Solar Analog HD 41248 on the Edge of a 7:5 Mean Motion Resonance. *The Astrophysical Journal*, 771:41.

Jewitt, D., Luu, J., and Marsden, B. G. (1992). 1992 QB1. *IAUcirc*, 5611.

Johnson, M. C., Cochran, W. D., Albrecht, S., Dodson-Robinson, S. E., Winn, J. N., and Gullikson, K. (2014). A Misaligned Prograde Orbit for Kepler-13 Ab via Doppler Tomography. *The Astrophysical Journal*, 790:30.

Jordi, C., Gebran, M., Carrasco, J. M., de Bruijne, J., Voss, H., Fabricius, C., Knude, J., Vallenari, A., Kohley, R., and Mora, A. (2010). Gaia broad band photometry. *Astronomy and Astrophysics*, 523:A48.

Jurić, M., Ivezić, Ž., Brooks, A., Lupton, R. H., Schlegel, D., Finkbeiner, D., Padmanabhan, N., Bond, N., Sesar, B., Rockosi, C. M., Knapp, G. R., Gunn, J. E., Sumi, T., Schneider, D. P., Barentine, J. C., Brewington, H. J., Brinkmann, J., Fukugita, M., Harvanek, M., Kleinman, S. J., Krzesinski, J., Long, D., Neilsen, Jr., E. H., Nitta, A., Snedden, S. A., and York, D. G. (2008). The Milky Way Tomography with SDSS. I. Stellar Number Density Distribution. *The Astrophysical Journal*, 673:864–914.

Kipping, D. M., Torres, G., Buchhave, L. A., Kenyon, S. J., Henze, C., Isaacson, H., Kolbl, R., Marcy, G. W., Bryson, S. T., Stassun, K., and Bastien, F. (2014). Discovery of a Transiting Planet near the Snow-line. *The Astrophysical Journal*, 795:25.

Kipping, D. M., Torres, G., Henze, C., Teachey, A., Isaacson, H., Petigura, E., Marcy, G. W., Buchhave, L. A., Chen, J., Bryson, S. T., and Sandford, E. (2016). A Transiting Jupiter Analog. *The Astrophysical Journal*, 820:112.

Kirk, B., Conroy, K., Prša, A., Abdul-Masih, M., Kochoska, A., Matijevič, G., Hambleton, K., Barclay, T., Bloemen, S., Boyajian, T., Doyle, L. R., Fulton, B. J., Hoekstra, A. J., Jek, K., Kane, S. R., Kostov, V., Latham, D., Mazeh, T., Orosz, J. A., Pepper, J., Quarles, B., Ragozzine, D., Shporer, A., Southworth, J., Stassun, K., Thompson, S. E., Welsh, W. F., Agol, E., Derekas, A., Devor, J., Fischer, D., Green, G., Gropp, J., Jacobs, T., Johnston, C., LaCourse, D. M., Saetre, K., Schwengeler, H., Toczyski, J., Werner, G., Garrett, M., Gore, J., Martinez, A. O., Spitzer, I., Stevick, J., Thomadis, P. C., Vrijmoet, E. H., Yenawine, M., Batalha, N., and Borucki, W. (2016). Kepler Eclipsing Binary Stars. VII. The Catalog of Eclipsing Binaries Found in the Entire Kepler Data Set. *The Astronomical Journal*, 151:68.

Kolodziejczak, J. J., Caldwell, D. A., Van Cleve, J. E., Clarke, B. D., Jenkins, J. M., Cote, M. T., Klaus, T. C., and Argabright, V. S. (2010). Flagging and correction of pattern noise in the Kepler focal plane array. In *High Energy, Optical, and Infrared Detectors for Astronomy IV*, volume 7742 of *Proceedings of the SPIE*, page 77421G.

Kroupa, P. (2001). On the variation of the initial mass function. *Monthly Notices of the Royal Astronomical Society*, 322:231–246.

Kupfer, T., Steeghs, D., Groot, P. J., Marsh, T. R., Nelemans, G., and Roelofs, G. H. A. (2016). UVES and X-Shooter spectroscopy of the emission line AM CVn systems GP Com and V396 Hya. *Monthly Notices of the Royal Astronomical Society*, 457:1828–1841.

LaCourse, D. M., Jek, K. J., Jacobs, T. L., Winarski, T., Boyajian, T. S., Rappaport, S. A., Sanchis-Ojeda, R., Conroy, K. E., Nelson, L., Barclay, T., Fischer, D. A., Schmitt, J. R., Wang, J., Stassun, K. G., Pepper, J., Coughlin, J. L., Shporer, A., and Prša, A. (2015). Kepler eclipsing binary stars - VI. Identification of eclipsing binaries in the K2 Campaign 0 data set. *Monthly Notices of the Royal Astronomical Society*, 452:3561–3592.

Lada, C. J. (2006). Stellar Multiplicity and the Initial Mass Function: Most Stars Are Single. *The Astrophysical Journal*, 640:L63–L66.

Lagrange, A.-M., Bonnefoy, M., Chauvin, G., Apai, D., Ehrenreich, D., Boccaletti, A., Gratadour, D., Rouan, D., Mouillet, D., Lacour, S., and Kasper, M. (2010). A Giant Planet Imaged in the Disk of the Young Star β Pictoris. *Science*, 329:57.

Lanza, A. F. (2010). Hot Jupiters and the evolution of stellar angular momentum. *Astronomy and Astrophysics*, 512:A77.

Larson, R. B. (1986). Bimodal star formation and remnant-dominated galactic models. *Monthly Notices of the Royal Astronomical Society*, 218:409–428.

Li, J., Tenenbaum, P., Twicken, J. D., Burke, C. J., Jenkins, J. M., Quintana, E. V., Rowe, J. F., and Seader, S. E. (2017). Kepler Data Processing Handbook: Data Validation II. Transit Model Fitting and Multiple Planet Search. Technical report, NASA.

Lillo-Box, J., Barrado, D., Moya, A., Montesinos, B., Montalbán, J., Bayo, A., Barbieri, M., Régulo, C., Mancini, L., Bouy, H., and Henning, T. (2014). Kepler-91b: a planet at the end of its life. Planet and giant host star properties via light-curve variations. *Astronomy and Astrophysics*, 562:A109.

Lin, D. N. C., Bodenheimer, P., and Richardson, D. C. (1996). Orbital migration of the planetary companion of 51 Pegasi to its present location. *Nature*, 380:606–607.

Lin, D. N. C. and Papaloizou, J. (1986). On the tidal interaction between protoplanets and the protoplanetary disk. III - Orbital migration of protoplanets. *Astrophysical Journal*, 309:846–857.

Lindegren, L., Lammers, U., Bastian, U., Hernández, J., Klioner, S., Hobbs, D., Bombrun, A., Michalik, D., Ramos-Lerate, M., Butkevich, A., Comoretto, G., Joliet, E., Holl, B., Hutton, A., Parsons, P., Steidelmüller, H., Abbas, U., Altmann, M., Andrei, A., Anton, S., Bach, N., Barache, C., Becciani, U., Berthier, J., Bianchi, L., Biermann, M., Bouquillon, S., Bourda, G., Brüsemeister, T., Bucciarelli, B., Busonero, D., Carlucci, T., Castañeda, J., Charlot, P., Clotet, M., Crosta, M., Davidson, M., de Felice, F., Drimmel, R., Fabricius, C., Fienga, A., Figueras, F., Fraile, E., Gai, M., Garralda, N., Geyer, R., González-Vidal, J. J., Guerra, R., Hambly, N. C., Hauser, M., Jordan, S., Lattanzi, M. G., Lenhardt, H., Liao, S., Löffler, W., McMillan, P. J., Mignard, F., Mora, A., Morbidelli, R., Portell, J., Riva, A., Sarasso, M., Serraller, I., Siddiqui, H., Smart, R., Spagna, A., Stampa, U., Steele, I., Taris, F., Torra, J., van Reeven, W., Vecchiato, A., Zschocke, S., de Bruijne, J., Gracia, G., Raison, F., Lister, T., Marchant, J., Messineo, R., Soffel, M., Osorio, J., de Torres, A., and O'Mullane, W. (2016). Gaia Data Release 1. Astrometry: one billion positions, two million proper motions and parallaxes. *Astronomy and Astrophysics*, 595:A4.

Liu, H.-G., Jiang, P., Huang, X., Yu, Z.-Y., Yang, M., Jia, M., Awiphan, S., Pan, X., Liu, B., Zhang, H., Wang, J., Li, Z., Du, F., Li, X., Lu, H., Zhang, Z., Tian, Q.-G., Li, B., Ji, T., Zhang, S., Shi, X., Wang, J., Zhou, J.-L., and Zhou, H. (2018). Searching for the Transit of the Earth-mass Exoplanet Proxima Centauri b in Antarctica: Preliminary Result. *The Astronomical Journal*, 155:12.

Lombardi, J. C., Thrall, A. P., Deneva, J. S., Fleming, S. W., and Grabowski, P. E. (2003). Modelling collision products of triple-star mergers. *Monthly Notices of the Royal Astronomical Society*, 345:762–780.

López-Morales, M., Haywood, R. D., Coughlin, J. L., Zeng, L., Buchhave, L. A., Giles, H. A. C., Affer, L., Bonomo, A. S., Charbonneau, D., Collier Cameron, A., Consentino, R., Dressing, C. D., Dumusque, X., Figueira, P., Fiorenzano, A. F. M., Harutyunyan, A., Johnson, J. A., Latham, D. W., Lopez, E. D., Lovis, C., Malavolta, L., Mayor, M., Micela, G., Molinari, E., Mortier, A., Motalebi, F., Nascimbeni, V., Pepe, F., Phillips, D. F., Piotto, G., Pollacco, D., Queloz, D., Rice, K., Sasselov, D., Segransan, D., Sozzetti, A., Udry, S., Vanderburg, A., and Watson, C. (2016). Kepler-21b: A Rocky Planet Around a V = 8.25 Magnitude Star. *The Astronomical Journal*, 152:204.

Malavolta, L., Borsato, L., Granata, V., Piotto, G., Lopez, E., Vanderburg, A., Figueira, P., Mortier, A., Nascimbeni, V., Affer, L., Bonomo, A. S., Bouchy, F., Buchhave, L. A., Charbonneau, D., Collier Cameron, A., Cosentino, R., Dressing, C. D., Dumusque, X., Fiorenzano, A. F. M., Harutyunyan, A., Haywood, R. D., Johnson, J. A., Latham, D. W., Lopez-Morales, M., Lovis, C., Mayor, M., Micela, G., Molinari, E., Motalebi, F., Pepe, F., Phillips, D. F., Pollacco, D., Queloz, D., Rice, K., Sasselov, D., Ségransan, D., Sozzetti, A., Udry, S., and Watson, C. (2017). The Kepler-19 System: A Thick-envelope Super-Earth with Two Neptune-mass Companions Characterized Using Radial Velocities and Transit Timing Variations. *The Astronomical Journal*, 153:224.

Mann, A. W., Gaidos, E., and Ansdell, M. (2013). Spectro-thermometry of M Dwarfs and Their Candidate Planets: Too Hot, Too Cool, or Just Right? *The Astrophysical Journal*, 779:188.

Marcos-Arenal, P., Zima, W., De Ridder, J., Aerts, C., Huygen, R., Samadi, R., Green, J., Piotto, G., Salmon, S., Catala, C., and Rauer, H. (2014). The PLATO Simulator: modelling of high-precision high-cadence space-based imaging. *Astronomy and Astrophysics*, 566:A92.

Marks, M. and Kroupa, P. (2012). Inverse dynamical population synthesis. Constraining the initial conditions of young stellar clusters by studying their binary populations. *Astronomy and Astrophysic*, 543:A8.

Martínez-Arnáiz, R., Maldonado, J., Montes, D., Eiroa, C., and Montesinos, B. (2010). Chromospheric activity and rotation of FGK stars in the solar vicinity. An estimation of the radial velocity jitter. *Astronomy and Astrophysics*, 520:A79.

Matijevič, G., Prša, A., Orosz, J. A., Welsh, W. F., Bloemen, S., and Barclay, T. (2012). Kepler Eclipsing Binary Stars. III. Classification of Kepler Eclipsing Binary Light Curves with Locally Linear Embedding. *The Astronomical Journal*, 143:123.

Maxted, P. F. L., Serenelli, A. M., and Southworth, J. (2015). Comparison of gyrochronological and isochronal age estimates for transiting exoplanet host stars. *Astronomy & Astrophysics*, 577:A90.

Mayor, M. and Queloz, D. (1995). A Jupiter-mass companion to a solar-type star. *Nature*, 378:355–359.

McCormac, J., Pollacco, D., Wheatley, P. J., West, R. G., Walker, S., Bento, J., Skillen, I., Faedi, F., Burleigh, M. R., Casewell, S. L., Chazelas, B., Genolet, L., Gibson, N. P., Goad, M. R., Lawrie, K. A., Ryans, R., Todd, I., Udry, S., and Watson, C. A. (2017). The Next Generation Transit Survey: Prototyping Phase. *Publications of the Astronomical Society of the Pacific*, 129(2):025002.

McCullough, P. R., Stys, J. E., Valenti, J. A., Fleming, S. W., Janes, K. A., and Heasley, J. N. (2005). The XO Project: Searching for Transiting Extrasolar Planet Candidates. *The Publications of the Astronomical Society of the Pacific*, 117:783–795.

Miglio, A., Chaplin, W. J., Farmer, R., Kolb, U., Girardi, L., Elsworth, Y., Appourchaux, T., and Handberg, R. (2014). Prospects for Detecting Asteroseismic Binaries in Kepler Data. *The Astrophysical Journal Letters*, 784:L3.

Mocquet, A., Grasset, O., and Sotin, C. (2014). Very high-density planets: a possible remnant of gas giants. *Philosophical Transactions of the Royal Society of London Series A*, 372:20130164–20130164.

Moe, M. and Di Stefano, R. (2017). Mind Your Ps and Qs: The Interrelation between Period (P) and Mass-ratio (Q) Distributions of Binary Stars. *The Astrophysical Journal Supplement Series*, 230:15.

Mordasini, C., Alibert, Y., Georgy, C., Dittkrist, K.-M., Klahr, H., and Henning, T. (2012). Characterization of exoplanets from their formation. II. The planetary mass-radius relationship. *Astronomy and Astrophysics*, 547:A112.

Morton, T. D., Bryson, S. T., Coughlin, J. L., Rowe, J. F., Ravichandran, G., Petigura, E. A., Haas, M. R., and Batalha, N. M. (2016). False Positive Probabilities for all Kepler Objects of Interest: 1284 Newly Validated Planets and 428 Likely False Positives. *The Astrophysical Journal*, 822:86.

Morton, T. D. and Johnson, J. A. (2011). On the Low False Positive Probabilities of Kepler Planet Candidates. *The Astrophysical Journal*, 738:170.

Muirhead, P. S., Vanderburg, A., Shporer, A., Becker, J., Swift, J. J., Lloyd, J. P., Fuller, J., Zhao, M., Hinkley, S., Pineda, J. S., Bottom, M., Howard, A. W., von Braun, K., Boyajian, T. S., Law, N., Baranec, C., Riddle, R., Ramaprakash, A. N., Tendulkar, S. P., Bui, K., Burse, M., Chordia, P., Das, H., Dekany, R., Punnadi, S., and Johnson, J. A. (2013). Characterizing the Cool KOIs. V. KOI-256: A Mutually Eclipsing Post-common Envelope Binary. *The Astrophysical Journal*, 767:111.

Mullally, F., Coughlin, J. L., Thompson, S. E., Rowe, J., Burke, C., Latham, D. W., Batalha, N. M., Bryson, S. T., Christiansen, J., Henze, C. E., Ofir, A., Quarles, B., Shporer, A., Van Eylen, V., Van Laerhoven, C., Shah, Y., Wolfgang, A., Chaplin, W. J., Xie, J.-W., Akeson, R., Argabright, V., Bachtell, E., Barclay, T., Borucki, W. J., Caldwell, D. A., Campbell,

J. R., Catanzarite, J. H., Cochran, W. D., Duren, R. M., Fleming, S. W., Fraquelli, D., Girouard, F. R., Haas, M. R., Hełminiak, K. G., Howell, S. B., Huber, D., Larson, K., Gautier, III, T. N., Jenkins, J. M., Li, J., Lissauer, J. J., McArthur, S., Miller, C., Morris, R. L., Patil-Sabale, A., Plavchan, P., Putnam, D., Quintana, E. V., Ramirez, S., Silva Aguirre, V., Seader, S., Smith, J. C., Steffen, J. H., Stewart, C., Stober, J., Still, M., Tenenbaum, P., Troeltzsch, J., Twicken, J. D., and Zamudio, K. A. (2015). Planetary Candidates Observed by Kepler. VI. Planet Sample from Q1–Q16 (47 Months). *The Astrophysical Journal Supplement Series*, 217:31.

Murphy, S. J., Bedding, T. R., and Shibahashi, H. (2016). A Planet in an 840 Day Orbit around a Kepler Main-sequence A Star Found from Phase Modulation of Its Pulsations. *The Astrophysical Journal Letters*, 827:L17.

Murray, N., Hansen, B., Holman, M., and Tremaine, S. (1998). Migrating Planets. *Science*, 279:69.

Muterspaugh, M. W., Lane, B. F., Kulkarni, S. R., Konacki, M., Burke, B. F., Colavita, M. M., Shao, M., Hartkopf, W. I., Boss, A. P., and Williamson, M. (2010). The Phases Differential Astrometry Data Archive. V. Candidate Substellar Companions to Binary Systems. *The Astronomical Journal*, 140:1657–1671.

O'Donovan, F. T., Charbonneau, D., Mandushev, G., Dunham, E. W., Latham, D. W., Torres, G., Sozzetti, A., Brown, T. M., Trauger, J. T., Belmonte, J. A., Rabus, M., Almenara, J. M., Alonso, R., Deeg, H. J., Esquerdo, G. A., Falco, E. E., Hillenbrand, L. A., Roussanova, A., Stefanik, R. P., and Winn, J. N. (2006). TrES-2: The First Transiting Planet in the Kepler Field. *The Astrophysical Journal*, 651:L61–L64.

Orosz, J. A., Welsh, W. F., Carter, J. A., Fabrycky, D. C., Cochran, W. D., Endl, M., Ford, E. B., Haghighipour, N., MacQueen, P. J., Mazeh, T., Sanchis-Ojeda, R., Short, D. R., Torres, G., Agol, E., Buchhave, L. A., Doyle, L. R., Isaacson, H., Lissauer, J. J., Marcy, G. W., Shporer, A., Windmiller, G., Barclay, T., Boss, A. P., Clarke, B. D., Fortney, J., Geary, J. C., Holman, M. J., Huber, D., Jenkins, J. M., Kinemuchi, K., Kruse, E., Ragozzine, D., Sasselov, D., Still, M., Tenenbaum, P., Uddin, K., Winn, J. N., Koch, D. G., and Borucki, W. J. (2012). Kepler-47: A Transiting Circumbinary Multiplanet System. *Science*, 337:1511.

Ortiz, M., Gandolfi, D., Reffert, S., Quirrenbach, A., Deeg, H. J., Karjalainen, R., Montañés-Rodríguez, P., Nespral, D., Nowak, G., Osorio, Y., and Palle, E. (2015). Kepler-432 b: a massive warm Jupiter in a 52-day eccentric orbit transiting a giant star. *Astronomy and Astrophysics*, 573:L6.

Pace, G. and Pasquini, L. (2004). The age-activity-rotation relationship in solar-type stars. *Astronomy and Astrophysics*, 426:1021–1034.

Patapis, P., Kühn, J., and Schmid, H. M. (2018). A simple optimized amplitude pupil mask for attempting to direct imaging of Proxima b with SPHERE/ZIMPOL at VLT. *ArXiv e-prints*.

Paxton, B., Bildsten, L., Dotter, A., Herwig, F., Lesaffre, P., and Timmes, F. (2011). Modules for Experiments in Stellar Astrophysics (MESA). *The Astrophysical Journal Supplement*, 192:3.

Paxton, B., Cantiello, M., Arras, P., Bildsten, L., Brown, E. F., Dotter, A., Mankovich, C., Montgomery, M. H., Stello, D., Timmes, F. X., and Townsend, R. (2013). Modules for Experiments in Stellar Astrophysics (MESA): Planets, Oscillations, Rotation, and Massive Stars. *The Astrophysical Journal Supplement*, 208:4.

Paxton, B., Marchant, P., Schwab, J., Bauer, E. B., Bildsten, L., Cantiello, M., Dessart, L., Farmer, R., Hu, H., Langer, N., Townsend, R. H. D., Townsley, D. M., and Timmes, F. X. (2015). Modules for Experiments in Stellar Astrophysics (MESA): Binaries, Pulsations, and Explosions. *The Astrophysical Journal Supplement Series*, 220:15.

Pepper, J., Pogge, R. W., DePoy, D. L., Marshall, J. L., Stanek, K. Z., Stutz, A. M., Poindexter, S., Siverd, R., O'Brien, T. P., Trueblood, M., and Trueblood, P. (2007). The Kilodegree Extremely Little Telescope (KELT): A Small Robotic Telescope for Large-Area Synoptic Surveys. *The Publications of the Astronomical Society of the Pacific*, 119:923–935.

Petigura, E. A., Howard, A. W., Marcy, G. W., Johnson, J. A., Isaacson, H., Cargile, P. A., Hebb, L., Fulton, B. J., Weiss, L. M., Morton, T. D., Winn, J. N., Rogers, L. A., Sinukoff, E., Hirsch, L. A., and Crossfield, I. J. M. (2017). The California-Kepler Survey. I. High Resolution Spectroscopy of 1305 Stars Hosting Kepler Transiting Planets. *ArXiv e-prints*.

Pollacco, D. L., Skillen, I., Collier Cameron, A., Christian, D. J., Hellier, C., Irwin, J., Lister, T. A., Street, R. A., West, R. G., Anderson, D. R., Clarkson, W. I., Deeg, H., Enoch, B., Evans, A., Fitzsimmons, A., Haswell, C. A., Hodgkin, S., Horne, K., Kane, S. R., Keenan, F. P., Maxted, P. F. L., Norton, A. J., Osborne, J., Parley, N. R., Ryans, R. S. I., Smalley, B., Wheatley, P. J., and Wilson, D. M. (2006). The WASP Project and the SuperWASP Cameras. *The Publications of the Astronomical Society of the Pacific*, 118:1407–1418.

Prša, A., Batalha, N., Slawson, R. W., Doyle, L. R., Welsh, W. F., Orosz, J. A., Seager, S., Rucker, M., Mjaseth, K., Engle, S. G., Conroy, K., Jenkins, J., Caldwell, D., Koch, D., and Borucki, W. (2011). Kepler Eclipsing Binary Stars. I. Catalog and Principal Characterization of 1879 Eclipsing Binaries in the First Data Release. *The Astronomical Journal*, 141:83.

Qian, S. B., Han, Z. T., Fernández Lajús, E., Zhu, L. Y., Li, L. J., Liao, W. P., and Zhao, E. G. (2015). Long-term Decrease and Cyclic Variation in the Orbital Period of the Eclipsing Dwarf Nova V2051 Oph. *The Astrophysical Journal Supplement Series*, 221:17.

Queloz, D., Bouchy, F., Moutou, C., Hatzes, A., Hébrard, G., Alonso, R., Auvergne, M., Baglin, A., Barbieri, M., Barge, P., Benz, W., Bordé, P., Deeg, H. J., Deleuil, M., Dvorak, R., Erikson, A., Ferraz Mello, S., Fridlund, M., Gandolfi, D., Gillon, M., Guenther, E., Guillot, T., Jorda, L., Hartmann, M., Lammer, H., Léger, A., Llebaria, A., Lovis, C., Magain, P., Mayor, M., Mazeh, T., Ollivier, M., Pätzold, M., Pepe, F., Rauer, H., Rouan, D., Schneider, J., Segransan, D., Udry, S., and Wuchterl, G. (2009). The CoRoT-7 planetary system: two orbiting super-Earths. *Astronomy and Astrophysics*, 506:303–319.

Raghavan, D., McAlister, H. A., Henry, T. J., Latham, D. W., Marcy, G. W., Mason, B. D., Gies, D. R., White, R. J., and ten Brummelaar, T. A. (2010). A Survey of Stellar Families: Multiplicity of Solar-type Stars. *The Astrophysical Journal Supplement*, 190:1–42.

Rauer, H., Catala, C., Aerts, C., Appourchaux, T., Benz, W., Brandeker, A., Christensen-Dalsgaard, J., Deleuil, M., Gizon, L., Goupil, M.-J., Güdel, M., Janot-Pacheco, E., Mas-Hesse, M., Pagano, I., Piotto, G., Pollacco, D., Santos, Ċ., Smith, A., Suárez, J.-C., Szabó,

R., Udry, S., Adibekyan, V., Alibert, Y., Almenara, J.-M., Amaro-Seoane, P., Eiff, M. A.-v., Asplund, M., Antonello, E., Barnes, S., Baudin, F., Belkacem, K., Bergemann, M., Bihain, G., Birch, A. C., Bonfils, X., Boisse, I., Bonomo, A. S., Borsa, F., Brandão, I. M., Brocato, E., Brun, S., Burleigh, M., Burston, R., Cabrera, J., Cassisi, S., Chaplin, W., Charpinet, S., Chiappini, C., Church, R. P., Csizmadia, S., Cunha, M., Damasso, M., Davies, M. B., Deeg, H. J., Díaz, R. F., Dreizler, S., Dreyer, C., Eggenberger, P., Ehrenreich, D., Eigmüller, P., Erikson, A., Farmer, R., Feltzing, S., de Oliveira Fialho, F., Figueira, P., Forveille, T., Fridlund, M., García, R. A., Giommi, P., Giuffrida, G., Godolt, M., Gomes da Silva, J., Granzer, T., Grenfell, J. L., Grotsch-Noels, A., Günther, E., Haswell, C. A., Hatzes, A. P., Hébrard, G., Hekker, S., Helled, R., Heng, K., Jenkins, J. M., Johansen, A., Khodachenko, M. L., Kislyakova, K. G., Kley, W., Kolb, U., Krivova, N., Kupka, F., Lammer, H., Lanza, A. F., Lebreton, Y., Magrin, D., Marcos-Arenal, P., Marrese, P. M., Marques, J. P., Martins, J., Mathis, S., Mathur, S., Messina, S., Miglio, A., Montalban, J., Montalto, M., Monteiro, M. J. P. F. G., Moradi, H., Moravveji, E., Mordasini, C., Morel, T., Mortier, A., Nascimbeni, V., Nelson, R. P., Nielsen, M. B., Noack, L., Norton, A. J., Ofir, A., Oshagh, M., Ouazzani, R.-M., Pápics, P., Parro, V. C., Petit, P., Plez, B., Poretti, E., Quirrenbach, A., Ragazzoni, R., Raimondo, G., Rainer, M., Reese, D. R., Redmer, R., Reffert, S., Rojas-Ayala, B., Roxburgh, I. W., Salmon, S., Santerne, A., Schneider, J., Schou, J., Schuh, S., Schunker, H., Silva-Valio, A., Silvotti, R., Skillen, I., Snellen, I., Sohl, F., Sousa, S. G., Sozzetti, A., Stello, D., Strassmeier, K. G., Švanda, M., Szabó, G. M., Tkachenko, A., Valencia, D., Van Grootel, V., Vauclair, S. D., Ventura, P., Wagner, F. W., Walton, N. A., Weingrill, J., Werner, S. C., Wheatley, P. J., and Zwintz, K. (2014). The PLATO 2.0 mission. *Experimental Astronomy*, 38:249–330.

Ricker, G. R., Winn, J. N., Vanderspek, R., Latham, D. W., Bakos, G. Á., Bean, J. L., Berta-Thompson, Z. K., Brown, T. M., Buchhave, L., Butler, N. R., Butler, R. P., Chaplin, W. J., Charbonneau, D., Christensen-Dalsgaard, J., Clampin, M., Deming, D., Doty, J., De Lee, N., Dressing, C., Dunham, E. W., Endl, M., Fressin, F., Ge, J., Henning, T., Holman, M. J., Howard, A. W., Ida, S., Jenkins, J. M., Jernigan, G., Johnson, J. A., Kaltenegger, L., Kawai, N., Kjeldsen, H., Laughlin, G., Levine, A. M., Lin, D., Lissauer, J. J., MacQueen, P., Marcy, G., McCullough, P. R., Morton, T. D., Narita, N., Paegert, M., Palle, E., Pepe, F., Pepper, J., Quirrenbach, A., Rinehart, S. A., Sasselov, D., Sato, B., Seager, S., Sozzetti, A., Stassun, K. G., Sullivan, P., Szentgyorgyi, A., Torres, G., Udry, S., and Villasenor, J. (2015). Transiting Exoplanet Survey Satellite (TESS). *Journal of Astronomical Telescopes, Instruments, and Systems*, 1(1):014003.

Robin, A. C., Luri, X., Reylé, C., Isasi, Y., Grux, E., Blanco-Cuaresma, S., Arenou, F., Babusiaux, C., Belcheva, M., Drimmel, R., Jordi, C., Krone-Martins, A., Masana, E., Mauduit, J. C., Mignard, F., Mowlavi, N., Rocca-Volmerange, B., Sartoretti, P., Slezak, E., and Sozzetti, A. (2012). Gaia Universe model snapshot. A statistical analysis of the expected contents of the Gaia catalogue. *Astronomy & Astrophysics*, 543:A100.

Robin, A. C., Reylé, C., Derrière, S., and Picaud, S. (2003). A synthetic view on structure and evolution of the Milky Way. *Astronomy and Astrophysics*, 409:523–540.

Rowe, J. F., Matthews, J. M., Seager, S., Sasselov, D., Kuschnig, R., Guenther, D. B., Moffat, A. F. J., Rucinski, S. M., Walker, G. A. H., and Weiss, W. W. (2009). Towards the Albedo of an Exoplanet: MOST Satellite Observations of Bright Transiting Exoplanetary Systems. In Pont, F., Sasselov, D., and Holman, M. J., editors, *Transiting Planets*, volume 253 of *IAU Symposium*, pages 121–127.

Rybizki, J., Demleitner, M., Fouesneau, M., Bailer-Jones, C., Rix, H.-W., and Andrae, R. (2018). A Gaia DR2 Mock Stellar Catalog. *ArXiv e-prints*.

Salpeter, E. E. (1955). The Luminosity Function and Stellar Evolution. *Astrophysical Journal*, 121:161.

Santerne, A., Hébrard, G., Deleuil, M., Havel, M., Correia, A. C. M., Almenara, J.-M., Alonso, R., Arnold, L., Barros, S. C. C., Behrend, R., Bernasconi, L., Boisse, I., Bonomo, A. S., Bouchy, F., Bruno, G., Damiani, C., Díaz, R. F., Gravallon, D., Guillot, T., Labrevoir, O., Montagnier, G., Moutou, C., Rinner, C., Santos, N. C., Abe, L., Audejean, M., Bendjoya, P., Gillier, C., Gregorio, J., Martinez, P., Michelet, J., Montaigut, R., Poncy, R., Rivet, J.-P., Rousseau, G., Roy, R., Suarez, O., Vanhuysse, M., and Verilhac, D. (2014). SOPHIE velocimetry of Kepler transit candidates. XII. KOI-1257 b: a highly eccentric three-month period transiting exoplanet. *Astronomy and Astrophysics*, 571:A37.

Seader, S., Jenkins, J. M., Tenenbaum, P., Twicken, J. D., Smith, J. C., Morris, R., Catanzarite, J., Clarke, B. D., Li, J., Cote, M. T., Burke, C. J., McCauliff, S., Girouard, F. R., Campbell, J. R., Kamal Uddin, A., Zamudio, K. A., Sabale, A., Henze, C. E., Thompson, S. E., and Klaus, T. C. (2015). Detection of Potential Transit Signals in 17 Quarters of Kepler Mission Data. *The Astrophysical Journal Supplement Series*, 217:18.

Shporer, A., Fuller, J., Isaacson, H., Hambleton, K., Thompson, S. E., Prša, A., Kurtz, D. W., Howard, A. W., and O'Leary, R. M. (2016). Radial Velocity Monitoring of Kepler Heartbeat Stars. *The Astrophysical Journal*, 829:34.

Slawson, R. W., Prša, A., Welsh, W. F., Orosz, J. A., Rucker, M., Batalha, N., Doyle, L. R., Engle, S. G., Conroy, K., Coughlin, J., Gregg, T. A., Fetherolf, T., Short, D. R., Windmiller, G., Fabrycky, D. C., Howell, S. B., Jenkins, J. M., Uddin, K., Mullally, F., Seader, S. E., Thompson, S. E., Sanderfer, D. T., Borucki, W., and Koch, D. (2011). Kepler Eclipsing Binary Stars. II. 2165 Eclipsing Binaries in the Second Data Release. *The Astronomical Journal*, 142:160.

Sokov, E. N., Vereshchagina, I. A., Gnedin, Y. N., Devyatkin, A. V., Gorshanov, D. L., Slesarenko, V. Y., Ivanov, A. V., Naumov, K. N., Zinov'ev, S. V., Bekhteva, A. S., Romas, E. S., Karashevich, S. V., and Kupriyanov, V. V. (2012). Observations of extrasolar planet transits with the automated telescopes of the Pulkovo Astronomical Observatory. *Astronomy Letters*, 38:180–190.

Southworth, J. (2008). Homogeneous studies of transiting extrasolar planets - I. Light-curve analyses. *Monthly Notices of the Royal Astronomical Society*, 386:1644–1666.

Southworth, J. (2010). Homogeneous studies of transiting extrasolar planets - III. Additional planets and stellar models. *Monthly Notices of the Royal Astronomical Society*, 408:1689–1713.

Southworth, J. (2011). Homogeneous studies of transiting extrasolar planets - IV. Thirty systems with space-based light curves. *Monthly Notices of the Royal Astronomical Society*, 417:2166–2196.

Southworth, J. (2013). The solar-type eclipsing binary system LL Aquarii. *Astronomy & Astrophysics*, 557:A119.

Southworth, J., Bruntt, H., and Buzasi, D. L. (2007). Eclipsing binaries observed with the WIRE satellite. II. β Aurigae and non-linear limb darkening in light curves. *Astronomy and Astrophysics*, 467:1215–1226.

Southworth, J., Hinse, T. C., Dominik, M., Glitrup, M., Jørgensen, U. G., Liebig, C., Mathiasen, M., Anderson, D. R., Bozza, V., Browne, P., Burgdorf, M., Calchi Novati, S., Dreizler, S., Finet, F., Harpsøe, K., Hessman, F., Hundertmark, M., Maier, G., Mancini, L., Maxted, P. F. L., Rahvar, S., Ricci, D., Scarpetta, G., Skottfelt, J., Snodgrass, C., Surdej, J., and Zimmer, F. (2009). Physical Properties of the 0.94-Day Period Transiting Planetary System WASP-18. *The Astrophysical Journal*, 707:167–172.

Southworth, J., Mancini, L., Madhusudhan, N., Mollière, P., Ciceri, S., and Henning, T. (2017). Detection of the Atmosphere of the 1.6 M $_\oplus$ Exoplanet GJ 1132 b. *The Astrophysical Journal*, 153:191.

Southworth, J., Maxted, P. F. L., and Smalley, B. (2004). Eclipsing binaries in open clusters - II. V453 Cyg in NGC 6871. *Monthly Notices of the Royal Astronomical Society*, 351:1277–1289.

Southworth, J., Smalley, B., Maxted, P. F. L., Claret, A., and Etzel, P. B. (2005). Absolute dimensions of detached eclipsing binaries - I. The metallic-lined system WW Aurigae. *Monthly Notices of the Royal Astronomical Society*, 363:529–542.

Staab, D., Haswell, C. A., Smith, G. D., Fossati, L., Barnes, J. R., Busuttil, R., and Jenkins, J. S. (2017). SALT observations of the chromospheric activity of transiting planet hosts: mass-loss and star-planet interactions. *Monthly Notices of the Royal Astronomical Society*, 466:738–748.

Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., Charbonneau, D., Deming, D., Dressing, C. D., Latham, D. W., Levine, A. M., McCullough, P. R., Morton, T., Ricker, G. R., Vanderspek, R., and Woods, D. (2015). The Transiting Exoplanet Survey Satellite: Simulations of Planet Detections and Astrophysical False Positives. *The Astrophysical Journal*, 809:77.

Tenenbaum, P., Christiansen, J. L., Jenkins, J. M., Rowe, J. F., Seader, S., Caldwell, D. A., Clarke, B. D., Li, J., Quintana, E. V., Smith, J. C., Stumpe, M. C., Thompson, S. E., Twicken, J. D., Van Cleve, J., Borucki, W. J., Cote, M. T., Haas, M. R., Sanderfer, D. T., Girouard, F. R., Klaus, T. C., Middour, C. K., Wohler, B., Batalha, N. M., Barclay, T., and Nickerson, J. E. (2012). Detection of Potential Transit Signals in the First Three Quarters of Kepler Mission Data. *The Astrophysical Journal Supplement*, 199:24.

Thompson, S. E., Coughlin, J. L., Hoffman, K., Mullally, F., Christiansen, J. L., Burke, C. J., Bryson, S., Batalha, N., Haas, M. R., Catanzarite, J., Rowe, J. F., Barentsen, G., Caldwell, D. A., Clarke, B. D., Jenkins, J. M., Li, J., Latham, D. W., Lissauer, J. J., Mathur, S., Morris, R. L., Seader, S. E., Smith, J. C., Klaus, T. C., Twicken, J. D., Wohler, B., Akeson, R., Ciardi, D. R., Cochran, W. D., Barclay, T., Campbell, J. R., Chaplin, W. J., Charbonneau, D., Henze, C. E., Howell, S. B., Huber, D., Prsa, A., Ramirez, S. V., Morton, T. D., Christensen-Dalsgaard, J., Dotson, J. L., Doyle, L., Dunham, E. W., Dupree, A. K., Ford, E. B., Geary, J. C., Girouard, F. R., Isaacson, H., Kjeldsen, H., Steffen, J. H., Quintana, E. V., Ragozzine, D., Shporer, A., Silva Aguirre, V., Still, M., Tenenbaum, P., Welsh, W. F., Wolfgang, A., Zamudio, K. A., Koch, D. G., and Borucki, W. J. (2017). Planetary Candidates Observed by Kepler. VIII. A Fully Automated Catalog With Measured Completeness and Reliability Based on Data Release 25. *ArXiv e-prints*.

Toonen, S., Hamers, A., and Portegies Zwart, S. (2016). The evolution of hierarchical triple star-systems. *Computational Astrophysics and Cosmology*, 3:6.

Traub, W. A. (2016). Kepler exoplanets: a new method of population analysis. *ArXiv e-prints*.

Uytterhoeven, K., Willems, B., Lefever, K., Aerts, C., Telting, J. H., and Kolb, U. (2004). Interpretation of the variability of the β Cephei star λ Scorpii. I. The multiple character. *Astronomy and Astrophysics*, 427:581–592.

Van Cleve, J. E. and Caldwell, D. A. (2016). Kepler Instrument Handbook. Technical report, NASA.

Van Eylen, V. and Albrecht, S. (2015). Eccentricity from Transit Photometry: Small Planets in Kepler Multi-planet Systems Have Low Eccentricities. *The Astrophysical Journal,*, 808:126.

Van Eylen, V., Lund, M. N., Silva Aguirre, V., Arentoft, T., Kjeldsen, H., Albrecht, S., Chaplin, W. J., Isaacson, H., Pedersen, M. G., Jessen-Hansen, J., Tingley, B., Christensen-Dalsgaard, J., Aerts, C., Campante, T. L., and Bryson, S. T. (2014). What Asteroseismology can do for Exoplanets: Kepler-410A b is a Small Neptune around a Bright Star, in an Eccentric Orbit Consistent with Low Obliquity. *The Astrophysical Journal*, 782:14.

van Hamme, W. (1993). New limb-darkening coefficients for modeling binary star light curves. *Astronomical Journal*, 106:2096–2117.

Wade, R. A. and Rucinski, S. M. (1985). Linear and quadratic limb-darkening coefficients for a large grid of LTE model atmospheres. *Astronomy and Astrophysics Supplement Series*, 60:471–484.

Wang, J., Fischer, D. A., Barclay, T., Picard, A., Ma, B., Bowler, B. P., Schmitt, J. R., Boyajian, T. S., Jek, K. J., LaCourse, D., Baranec, C., Riddle, R., Law, N. M., Lintott, C., Schawinski, K., Simister, D. J., Grégoire, B., Babin, S. P., Poile, T., Jacobs, T. L., Jebson, T., Omohundro, M. R., Schwengeler, H. M., Sejpka, J., Terentev, I. A., Gagliano, R., Paakkonen, J.-P., Otnes Berge, H. K., Winarski, T., Green, G. R., Schmitt, A. R., Kristiansen, M. H., and Hoekstra, A. (2015). Planet Hunters. VIII. Characterization of 41 Long-period Exoplanet Candidates from Kepler Archival Data. *The Astrophysical Journal*, 815:127.

Ward, W. R. (1997). Protoplanet Migration by Nebula Tides. *Icarus*, 126:261–281.

Weidenschilling, S. J. and Marzari, F. (1996). Gravitational scattering as a possible origin for giant planets at small stellar distances. *Nature*, 384:619–621.

Wheatley, P. J., West, R. G., Goad, M. R., Jenkins, J. S., Pollacco, D. L., Queloz, D., Rauer, H., Udry, S., Watson, C. A., Chazelas, B., Eigmüller, P., Lambert, G., Genolet, L., McCormac, J., Walker, S., Armstrong, D. J., Bayliss, D., Bento, J., Bouchy, F., Burleigh, M. R., Cabrera, J., Casewell, S. L., Chaushev, A., Chote, P., Csizmadia, S., Erikson, A., Faedi, F., Foxell, E., Gänsicke, B. T., Gillen, E., Grange, A., Günther, M. N., Hodgkin, S. T., Jackman, J., Jordán, A., Louden, T., Metrailler, L., Moyano, M., Nielsen, L. D., Osborn, H. P., Poppenhaeger, K., Raddi, R., Raynard, L., Smith, A. M. S., Soto, M., and Titz-Weider, R. (2018). The Next Generation Transit Survey (NGTS). *Monthly Notices of the Royal Astronomical Society*, 475:4476–4493.

Willems, B. and Kolb, U. (2002). Population synthesis of wide binary millisecond pulsars. *Monthly Notice of the Royal Astronomical Society*, 337:1004–1016.

Willems, B. and Kolb, U. (2003). On the detection of pre-low-mass X-ray binaries. *Monthly Notice of the Royal Astronomical Society*, 343:949–958.

Willems, B. and Kolb, U. (2004). Detached white dwarf main-sequence star binaries. *Astronomy and Astrophysics*, 419:1057–1076.

Willems, B., Kolb, U., and Justham, S. (2006). Eclipsing binaries in extrasolar planet transit surveys: the case of SuperWASP. *Monthly Notices of the Royal Astronomical Society*, 367:1103–1112.

Willems, B., Kolb, U., Sandquist, E. L., Taam, R. E., and Dubus, G. (2005). Angular Momentum Losses and the Orbital Period Distribution of Cataclysmic Variables below the Period Gap: Effects of Circumbinary Disks. *The Astrophysical Journal*, 635:1263–1280.

Willems, B., Taam, R. E., Kolb, U., Dubus, G., and Sandquist, E. L. (2007). Theoretical Orbital Period Distributions of Cataclysmic Variables above the Period Gap: Effects of Circumbinary Disks. *The Astrophysical Journal*, 657:465–481.

Wolszczan, A. and Frail, D. A. (1992). A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355:145–147.

Wu, Y. and Lithwick, Y. (2011). Secular Chaos and the Production of Hot Jupiters. *The Astrophysical Journal*, 735:109.

Zoccali, M., Renzini, A., Ortolani, S., Greggio, L., Saviane, I., Cassisi, S., Rejkuba, M., Barbuy, B., Rich, R. M., and Bica, E. (2003). Age and metallicity distribution of the Galactic bulge from extensive optical and near-IR stellar photometry. *Astronomy and Astrophysics*, 399:931–956.

Appendix A

Exoplanets: "Other" discovery methods

Planets in this category on Exoplanet.eu (19 March 2018) are:

BD+20 1790: Analysis of stellar activity on a young, very active K5Ve star (Hernán-Obispo et al., 2015).

Kepler-19 c: Detection of an additional non-transiting planet through transit timing variations (Malavolta et al., 2017).

HD 41248 c: Applying a mean motion resonance algorithm to archival spectroscopic data (Jenkins et al., 2013).

GJ 3293 c: In the discovery paper for this system, Astudillo-Defru et al. (2015) refer to two Neptunes (b and d) in 4:1 resonance, detected through radial velocity (RV) measurements, and a more tentative detection of a superEarth, c.

V2051 Oph b: Qian et al. (2015) derive the presence of a planet with $M \sin i = 7.3 \pm 0.7$ M_J with a semi-major axis of 9.0 AU and an eccentricity of 0.37, as an explanation for the rate of period decrease observed in this cataclysmic variable (CV) binary system.

KIC 7917485 b: Murphy et al. (2016) detected a 12 M_J planet with a period of 840 ± 20 days and an eccentricity of 0.15 around a class A star through pulsational phase shifts.

Kepler-20 g: Buchhave et al. (2016) discovered a sixth, non-transiting planet while taking RV measurements of the *Kepler*-20 system, already known to have five transiting planets.

HD 163296 b and HD 163296 c: Isella et al. (2016) infer the presence of two Saturn-mass planets in the protoplanetary disc around this star at distances of 100 and 160 AU based on depletion of dust and CO.

DW Uma b: In their studies of the eclipsing binary system, Boyd et al. (2017) derive the presence of a possible third body with $M = 10.06$ M_J and a semi-major axis of 5.80 AU through TTVs and conclude that this body, if it exists, is more likely to be a planet than a star.

OGLE-GD-ECL-11388 b: Hong et al. (2017) infer the presence of a body with 12.5 M_J, on the boundary between a planet and a brown dwarf, from transit timing variations (TTVs) noted during their observations of this binary system.

GP Com b and V396 Hya b: Exoplanet.eu states that Kupfer et al. (2016) detected an object through Doppler tomography in each of these systems. However, Kupfer et al. (2016) state that both these systems are AM CVn binaries, with periods of 46.57 minutes and about 65.1 minutes respectively. The masses of the objects are quoted in M_J, but it is unlikely that these bodies are planets, and Kupfer et al. (2016) do not make that claim.

KIC 5095269 (AB) b: A planetary mass body was found while searching for eclipsing binaries in the Kepler archive (Getley et al., 2017).

Appendix B

Kepler circumbinary planets (S-type)

Confirmed *Kepler* planets known to be circumbinary in an S-type orbit (ie orbiting just one star in the system) (14 September 2017) are:

Kepler-13 A b (Johnson et al., 2014). P = 1.8 days, R = 21.42 $R_\oplus$. Through RV measurements, the authors confirm that this planet is in a misaligned prograde orbit, and that the star it is orbiting is part of a triple system: A has a photometric companion with two spectroscopic components, B and C.

Kepler-21 b (López-Morales et al., 2016). P = 2.8 days, R = 1.69 $R_\oplus$. The authors believe this is a rocky planet whose atmosphere has evaporated. At V = 8.25 and F6IV, the host star is bright enough to fall in the Plato P1 population.

Kepler-68 A b and A c (Gilliland et al., 2013). (b) P = 5.4 days, R = 2.21 $R_\oplus$, (c) P = 9.6 days, R = 0.91 $R_\oplus$. RV follow-up indicates a third Jovian-mass non-transiting planet orbiting beyond the two transiting planets, with a period of 580 $\pm$ 15 days.

Kepler-296 b, c, d, e and f (Barclay et al., 2015). (b) P = 10.9 days, R = 1.14 $R_\oplus$. (c) P = 5.8 days, R = 1.38 $R_\oplus$. (d) P = 19.9 days, R = 1.52 $R_\oplus$. (e) P = 34.1 days, R = 1.18 $R_\oplus$. (f) P = 63.3 days, R = 1.06 $R_\oplus$. MCMC modelling indicates that the five planets orbit the primary in this binary system, consisting of two M-class stars separated by 0.2".

Kepler-420 A b (Santerne et al., 2014). P = 86.6 days, R = 7.62 $R_\oplus$. Confirmed through RV follow up, this planet orbits the primary in a metal-rich, relatively old binary system. The planet has a highly eccentric orbit of 0.772 $\pm$ 0.045.

Kepler-432 b (Ortiz et al., 2015). P = 52.5 days, R = 12.39 $R_\oplus$. This planet orbits an evolved star on the red giant branch: it is the third body in the system based on RV measurements. It has an eccentric orbit of 0.478 $\pm$ 0.004.

Kepler-410 A b (Van Eylen et al., 2014). P = 17.8 days, R = 2.43 $R_\oplus$. This planet was confirmed through asteroseismology to be orbiting the primary star in a binary star system, with an eccentricity of $0.17^{+0.07}_{-0.06}$.

Kepler-444 b, c, d, e and f (Van Eylen and Albrecht, 2015). (b) P = 3.6 days, R = 0.5 $R_\oplus$. (c) P = 4.5 days, R = 0.65 $R_\oplus$. (d) P = 6.2 days, R = 0.65 $R_\oplus$. (e) P = 7.7 days, R = 0.62 $R_\oplus$. (f) P = 9.7 days, R = 0.95 $R_\oplus$. These planets were included in a study of planets with low eccentricity: 0.08, 0.12, 0.18, 0.02 and 0.58 respectively.

Kepler-449 b and c (Van Eylen and Albrecht, 2015). (b) P = 12.6 days, R = 1.53 $R_\oplus$. (c) P = 33.7 days, R = 1.86 $R_\oplus$. These planet were included in a study of planets with low eccentricity: 0.03 and 0.05 respectively.

Kepler-450 b, c and d (Van Eylen and Albrecht, 2015). (b) P = 28.5 days, R = 6.08 $R_\oplus$. (c) P = 15.4 days, R = 2.59 $R_\oplus$. (d) P = 7.5 days, R = 0.96 $R_\oplus$. These planets were included in a study of planets with low eccentricity: 0.02, 0.02 and 0.14 respectively.

Appendix C

Detailed results of exoplanet simulations

Table C.1 Intrinsic exoplanet distribution, single long cadence ('zeta') (red) and benchmark estimate ('omega') (black) rms photometric precision. Data is given as *log N* in the *log R log P* plane. Uncertainty is $\sqrt{N}$.

Period/days (log)	*Radius*/$R_\oplus$ (log) **-.4 - -.2**	**-.2 - 0.0**	**0.0 - 0.2**	**0.2 - 0.4**	**0.4 - 0.6**	**0.6 - 0.8**	**0.8 - 1.0**	**1.0 - 1.2**	**1.2 - 1.4**	**1.4 - 1.6**
-0.4 - -0.2	-4.95 -6.10	-2.86 -3.01	-3.70 -3.69	-4.50 -4.29	-7.84 -8.23	-8.01 -9.75	-8.09 -9.44	-8.28 -9.73	-8.31 -9.58	-8.32 8.26
-0.2 - 0.0	-1.83 -2.19	-2.23 -2.41	-2.58 -2.58	-4.04 -3.73	-5.24 -4.58	-8.05 -9.64	-8.10 -9.51	-19.80 -39.95	-5.72 -4.64	-8.17 -9.63
0.0 - 0.2	-4.48 -7.44	-2.36 -2.55	-2.59 -2.59	-3.55 -3.24	-4.96 -4.28	-5.47 -4.69	-19.60 -41.83	-5.37 -4.36	-19.72 -41.47	-8.17 -15.37
0.2 - 0.4	-4.21 -5.67	-1.52 -1.71	-2.20 -2.19	-3.05 -2.73	-4.20 -3.53	-4.79 -3.95	-5.03 -4.05	-4.62 -3.56	-5.20 -4.07	-8.04 -9.23
0.4 - 0.6	-1.71 -2.10	-1.42 -1.61	-1.74 -1.74	-2.56 -2.25	-3.61 -2.95	-4.19 -3.34	-4.64 -3.72	-4.12 -3.12	-5.42 -4.17	-7.89 -9.34
0.6 - 0.8	-1.74 -2.06	-1.20 -1.37	-1.63 -1.63	-2.22 -1.90	-3.01 -2.35	-4.08 -3.22	-4.37 -3.48	-4.40 -3.47	-4.85 -3.85	-12.54 -40.28
0.8 - 1.0	-0.96 -1.36	-1.36 -1.56	-1.48 -1.47	-1.95 -1.64	-2.77 -2.11	-3.62 -2.77	-4.11 -3.27	-4.44 -3.64	-7.57 -9.13	-5.33 -4.37
1.0 - 1.2	-1.04 -1.41	-1.51 -1.69	-1.42 -1.42	-1.78 -1.46	-2.55 -1.89	-3.65 -2.79	-4.18 -3.35	-4.40 -3.60	-4.77 -3.92	-7.65 -9.06
1.2 - 1.4	-3.57 -6.42	-1.46 -1.63	-1.49 -1.49	-1.74 -1.43	-2.29 -1.63	-3.33 -2.47	-3.88 -3.07	-4.10 -3.33	-7.21 -8.88	-7.39 -8.98
1.4 - 1.6	-1.11 -1.41	-1.33 -1.55	-1.85 -1.88	-1.79 -1.47	-2.31 -1.66	-3.39 -2.53	-3.99 -3.18	-4.09 -3.30	-7.62 -8.94	-4.94 -4.09
1.6 - 1.8	-4.30 -6.19	-4.68 -7.18	-2.06 -2.14	-1.76 -1.43	-2.29 -1.62	-3.32 -2.47	-3.57 -2.72	-3.96 -3.14	-7.03 -8.81	-11.92 -39.70
1.8 - 2.0	-4.21 -6.00	-3.73 -5.56	-2.00 -2.03	-2.06 -1.74	-2.39 -1.73	-3.02 -2.17	-3.55 -2.69	-4.65 -3.87	-6.93 -8.67	-4.57 -3.75
2.0 - 2.2	-3.55 -4.48	-4.28 -5.42	-1.92 -1.93	-2.11 -1.78	-2.50 -1.83	-3.22 -2.39	-3.58 -2.74	-3.37 -2.60	-6.84 -8.39	-6.85 -8.31
2.2 - 2.4	-4.20 -4.42	-4.07 -6.83	-2.58 -2.76	-2.24 -1.91	-2.54 -1.89	-3.23 -2.42	-3.24 -2.38	-3.96 -3.10	-6.81 -8.35	-6.91 -7.06
2.4 - 2.6	-4.07 -5.87	-3.94 -6.45	-1.98 -2.08	-3.01 -2.72	-3.10 -2.40	-3.38 -2.54	-3.40 -2.63	-3.54 -2.76	-6.62 -8.30	-6.76 -8.19
2.6 - 2.8	-4.21 -5.62	-3.61 -6.51	-4.75 -12.80	-5.44 -7.49	-2.87 -2.18	-3.76 -2.97	-3.58 -2.66	-6.51 -7.96	-6.53 -8.09	-6.50 -8.15
2.8 - 3.0	-3.73 -5.52	-3.52 -6.25	-4.27 -7.08	-5.25 -7.54	-5.73 -7.52	-3.60 -2.81	-8.53 -38.51	-6.29 -7.92	-6.34 -7.73	-6.52 -6.66
3.0 - 3.2	-3.90 -3.94	-3.60 -4.64	-4.49 -6.79	-5.15 -7.27	-5.69 -7.29	-15.37 -38.54	-3.65 -2.67	-6.20 -13.68	-6.21 -7.88	-6.24 -6.53

Table C.2 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, $V < 26$. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	$10^{-0.4}$-$10^{-0.2}$ $R_\oplus$	$10^{-0.2}$-10^{0} $R_\oplus$	10^{0}-$10^{.2}$ $R_\oplus$	$10^{.2}$-$10^{.4}$ $R_\oplus$	$10^{.4}$-$10^{.6}$ $R_\oplus$	$10^{.6}$-$10^{.8}$ $R_\oplus$	$10^{.8}$-10^{1} $R_\oplus$	10^{1}-$10^{1.2}$ $R_\oplus$	$10^{1.2}$-$10^{1.4}$ $R_\oplus$
Observed *Kepler* DR25	10	126	490	797	584	133	67	68	9
$V < 26$									
32 unblended									
LLN 'zeta'	146 ±10	1650 ±40	6120 ±90	9300 ±80	6830 ± 90	1960 ± 40	1460 ±30	2220 ±50	420 ±30
LLN 'omega'	7 ± 2	116 ± 3	680 ±20	2060 ±40	3350 ± 40	1470 ± 20	1130 ±20	1890 ±30	503 ± 8
LLS 'zeta'	160 ±10	1840 ±50	6970 ±70	10120 ±70	6840 ± 50	1800 ± 60	1190 ±20	1530 ±20	240 ±10
LLS 'omega'	10 ± 3	129 ± 9	763 ± 7	2280 ±20	3390 ± 60	1360 ± 20	910 ±20	1290 ±10	290 ±10
32 inc blends									
LLN 'zeta'	150 ±10	1700 ±40	5650 ±90	8500 ±60	5490 ± 90	1540 ± 30	1110 ±20	800 ±30	90 ±10
LLN 'omega'	7 ± 2	127 ± 5	640 ±10	1970 ±50	2710 ± 30	1170 ± 20	880 ±10	710 ±20	107 ± 5
LLS 'zeta'	160 ±10	1900 ±50	6520 ±70	9450 ±60	5770 ± 40	1510 ± 50	970 ±30	740 ±10	80 ±10
LLS 'omega'	10 ± 2	138 ± 9	730 ±10	2210 ±30	2880 ± 40	1200 ±100	780 ±20	628 ±10	93 ± 4
28 unblended									
LLN 'zeta'	119 ± 9	1410 ±30	5400 ±70	8420 ±90	6300 ±100	1850 ± 30	1400 ±30	2130 ±50	400 ±20
LLN 'omega'	6 ± 6	100 ± 4	600 ±20	1870 ±30	3100 ± 40	1400 ± 20	1080 ±20	1820 ±30	484 ± 4
LLS 'zeta'	134 ±10	1560 ±40	6140 ±60	9200 ±60	6380 ± 60	1720 ± 60	1150 ±20	1490 ±20	230 ±10
LLS 'omega'	9 ± 3	109 ± 9	680 ±10	2060 ±30	3160 ± 50	1300 ± 20	880 ±20	1259 ±10	277 ± 9
28 inc blends									
LLN 'zeta'	118 ± 9	1440 ±30	4990 ±70	7720 ±60	5130 ± 80	1480 ± 30	1080 ±20	790 ±30	90 ±10
LLN 'omega'	6 ± 2	108 ± 5	570 ±10	1790 ±40	2532 ±20	1120 ± 20	860 ±10	700 ±20	108 ± 5
LLS 'zeta'	130 ±10	1610 ±50	5740 ±70	8630 ±50	5410 ± 40	1450 ± 50	960 ±30	740 ±10	80 ±10
LLS 'omega'	9 ± 3	116 ± 9	650 ±10	2010 ±40	2700 ± 40	1100 ± 30	768 ±20	630 ±10	93 ± 4
24 unblended									
LLN 'zeta'	93 ± 7	1170 ±40	4660 ±70	7480 ±80	5800 ± 90	1750 ± 30	1340 ±30	2040 ±50	380 ±20
LLN 'omega'	4 ± 1	84 ± 6	520 ±20	1660 ±30	2840 ± 30	1310 ± 20	1030 ±20	1740 ±30	460 ± 6
LLS 'zeta'	110 ±10	1290 ±40	5310 ±60	8220 ±50	5880 ± 50	1630 ± 60	1110 ±20	1450 ±20	220 ±10
LLS 'omega'	7 ± 2	90 ± 7	586 ±10	1840 ±30	2910 ± 40	1230 ± 20	850 ±20	1220 ±10	270 ±10
24 inc blends									
LLN 'zeta'	91 ± 8	1200 ±40	4310 ±60	6890 ±60	4730 ± 70	1400 ± 30	1040 ±20	790 ±30	90 ±10
LLN 'omega'	4 ± 1	90 ± 6	490 ±10	1590 ±40	2330 ± 20	1060 ± 10	820 ±10	700 ±20	107 ± 5
LLS 'zeta'	110 ±10	1330 ±40	4980 ±70	7730 ±60	5010 ± 40	1390 ± 50	930 ±30	730 ±10	80 ±10
LLS 'omega'	7 ± 2	98 ± 7	560 ±20	1790 ±30	2490 ± 40	1050 ± 20	740 ±20	620 ±10	93 ± 4

Table C.3 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, $V < 16$, F5-M. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	$10^{-0.4}$-$10^{-0.2}$ $R_\oplus$	$10^{-0.2}$-10^{0} $R_\oplus$	10^{0}-$10^{.2}$ $R_\oplus$	$10^{.2}$-$10^{.4}$ $R_\oplus$	$10^{.4}$-$10^{.6}$ $R_\oplus$	$10^{.6}$-$10^{.8}$ $R_\oplus$	$10^{.8}$-10^{1} $R_\oplus$	10^{1}-$10^{1.2}$ $R_\oplus$	$10^{1.2}$-$10^{1.4}$ $R_\oplus$
Observed *Kepler* DR25	10	124	488	791	578	126	65	66	7
$V < 16$, F5-M									
32 unblended									
LLN 'zeta'	140 ±10	1530 ±40	5350 ±100	7700 ±60	5260 ±90	1250 ±20	690 ±20	650 ±20	70 ±6
LLN 'omega'	6 ± 2	108 ± 3	600 ± 20	1710 ±40	2560 ±40	950 ±10	537 ± 9	560 ±30	86 ±6
LLS 'zeta'	160 ±10	1700 ±40	6140 ± 80	8490 ±70	5390 ±50	1220 ±40	630 ±10	590 ±20	60 ±8
LLS 'omega'	10 ± 3	120 ± 8	670 ± 5	1900 ±20	2650 ±30	932 ±20	480 ±20	498 ± 9	74 ±2
32 inc blends									
LLN 'zeta'	140 ±10	1560 ±40	4950 ± 90	7090 ±40	4350 ±80	1080 ±20	660 ±20	470 ±20	53 ±6
LLN 'omega'	7 ± 2	118 ± 5	560 ± 20	1640 ±50	2150 ±30	825 ± 8	520 ±10	410 ±20	65 ±4
LLS 'zeta'	150 ±10	1750 ±50	5750 ± 80	7990 ±70	4660 ±40	1100 ±40	620 ±20	470 ±20	50 ±6
LLS 'omega'	10 ± 2	128 ± 8	643 ± 6	1851 ±30	2300 ±20	840 ±20	480 ±20	402 ± 9	59 ±2
28 unblended									
LLN 'zeta'	114 ± 9	1310 ±30	4750 ± 80	6990 ±60	4920 ±90	1210 ±20	680 ±20	650 ±20	70 ±6
LLN 'omega'	5 ± 2	93 ± 4	530 ± 20	1560 ±30	2390 ±50	920 ±10	530 ±10	560 ±30	86 ±6
LLS 'zeta'	128 ± 9	1450 ±40	5430 ± 80	7750 ±50	5060 ±50	1180 ±40	624 ±10	590 ±20	60 ±8
LLS 'omega'	8 ± 3	101 ± 8	598 ± 5	1730 ±20	2490 ±30	900 ±20	480 ±20	498 ± 9	74 ±2
28 inc blends									
LLN 'zeta'	113 ± 9	1340 ±30	4400 ± 80	6470 ±50	4080 ±80	1050 ±20	650 ±20	470 ±20	53 ±6
LLN 'omega'	5 ± 2	101 ± 5	500 ± 20	1500 ±40	2010 ±30	800 ±10	510 ±10	410 ±20	65 ±4
LLS 'zeta'	126 ± 9	1490 ±50	5100 ± 70	7320 ±60	4390 ±40	1060 ±40	610 ±20	470 ±20	50 ±6
LLS 'omega'	9 ± 3	107 ± 7	574 ± 9	1690 ±30	2160 ±20	817 ±20	470 ±20	401 ± 9	59 ±2
24 unblended									
LLN 'zeta'	90 ± 7	1080 ±40	4120 ± 70	6240 ±60	4540 ±90	1160 ±20	670 ±20	650 ±20	70 ±6
LLN 'omega'	4 ±1	79 ± 6	460 ± 20	1380 ±30	2210 ±30	880 ±10	520 ±10	550 ±30	86 ±6
LLS 'zeta'	100 ±10	1190 ±40	4720 ± 70	6950 ±60	4690 ±50	1140 ±40	620 ±10	590 ±20	60 ±8
LLS 'omega'	7 ± 2	85 ± 6	517 ± 7	1540 ±20	2300 ±20	870 ±20	470 ±10	497 ±10	74 ±2
24 inc blends									
LLN 'zeta'	87 ± 8	1110 ±40	3820 ± 60	5790 ±50	3780 ±70	1010 ±20	640 ±20	470 ±20	53 ±6
LLN 'omega'	4 ± 2	85 ± 6	430 ± 20	1340 ±40	1860 ±30	760 ± 9	500 ±10	410 ±20	65 ±4
LLS 'zeta'	102 ± 9	1240 ±50	4430 ± 70	6580 ±70	4080 ±30	1030 ±40	600 ±20	470 ±20	50 ±6
LLS 'omega'	7 ± 2	90 ± 5	500 ± 10	1520 ±30	2010 ±20	790 ±20	470 ±20	400 ± 8	59 ±2

Table C.4 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, P1. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	$10^{-0.4}$-$10^{-0.2}$ $R_\oplus$	$10^{-0.2}$-10^{0} $R_\oplus$	10^{0}-$10^{.2}$ $R_\oplus$	$10^{.2}$-$10^{.4}$ $R_\oplus$	$10^{.4}$-$10^{.6}$ $R_\oplus$	$10^{.6}$-$10^{.8}$ $R_\oplus$	$10^{.8}$-10^{1} $R_\oplus$	10^{1}-$10^{1.2}$ $R_\oplus$	$10^{1.2}$-$10^{1.4}$ $R_\oplus$
Observed *Kepler* DR25	2	6	7	7	4	2	0	0	1
P1									
32 unblended									
LLN 'zeta'	80 ±10	550 ±20	810 ±40	384 ±20	100 ± 6	12 ±4	6 ±3	5 ±3	0.3 ±0.5
LLN 'omega'	4 ± 2	39 ± 6	84 ± 4	85 ± 9	47 ± 4	8 ±1	3.2 ±0.7	3 ±2	0.2 ±0.2
LLS 'zeta'	87 ± 7	610 ±20	980 ±30	460 ±30	110 ±10	13 ±4	5 ±2	5 ±1	0.6 ±0.7
LLS 'omega'	5 ± 1	46 ± 5	107 ± 5	99 ± 8	54 ± 5	10 ±2	4.7 ±0.7	5 ±2	0.7 ±0.4
32 inc blends									
LLN 'zeta'	80 ±10	570 ±20	780 ±30	390 ±20	95 ± 7	12 ±4	6 ±3	4 ±2	0.3 ±0.5
LLN 'omega'	4 ± 2	40 ± 5	81 ± 4	86 ± 8	45 ± 4	8 ±1	3.3 ±0.7	3 ±2	0.2 ±0.2
LLS 'zeta'	86 ± 7	640 ±20	940 ±40	460 ±30	110 ±10	13 ±4	5 ±2	5 ±1	0.6 ±0.7
LLS 'omega'	5 ± 1	49 ± 5	104 ± 4	101 ± 7	52 ± 5	10 ±2	5.0 ±0.8	4 ±2	0.7 ±0.4
28 unblended									
LLN 'zeta'	65 ± 7	494 ±20	770 ±40	378 ±20	99 ± 6	12 ±4	6 ±3	5 ±3	0.3 ±0.5
LLN 'omega'	4 ± 2	33 ± 4	78 ± 4	83 ± 7	47 ± 4	8 ±2	3.2 ±0.7	3 ±2	0.2 ±0.2
LLS 'zeta'	72 ± 5	540 ±20	930 ±30	450 ±30	110 ±10	13 ±4	5 ±2	5 ±1	0.6 ±0.7
LLS 'omega'	4 ± 2	41 ± 4	103 ± 4	97 ± 8	54 ± 5	10 ±2	4.7 ±0.7	5 ±2	0.7 ±0.4
28 inc blends									
LLN 'zeta'	65 ± 7	510 ±20	740 ±30	380 ±20	94 ± 8	12 ±4	6 ±3	4 ±2	0.3 ±0.5
LLN 'omega'	4 ± 2	35 ± 5	76 ± 5	85 ± 7	45 ± 4	8 ±1	3.3 ±0.7	3 ±2	0.2 ±0.2
LLS 'zeta'	71 ± 5	560 ±20	900 ±30	450 ±30	110 ±10	13 ±4	5 ±2	5 ±1	0.6 ±0.7
LLS 'omega'	4 ± 2	44 ± 5	100 ± 4	99 ± 7	51 ± 5	10 ±2	5.0 ±0.8	4 ±2	0.7 ±0.4
24 unblended									
LLN 'zeta'	52 ± 7	430 ±20	730 ±30	370 ±20	99 ± 6	12 ±4	6 ±3	5 ±3	0.3 ±0.5
LLN 'omega'	3 ± 1	30 ± 4	74 ± 3	81 ± 7	47 ± 4	8 ±1	3.2 ±0.7	3 ±2	0.2 ±0.2
LLS 'zeta'	58 ± 6	470 ±20	870 ±30	440 ±30	110 ±10	13 ±4	5 ±2	5 ±1	0.6 ±0.4
LLS 'omega'	3 ± 1	37 ± 4	97 ± 4	95 ± 9	53 ± 5	10 ±2	4.7 ±0.7	5 ±2	0.7 ±0.4
24 inc blends									
LLN 'zeta'	52 ± 7	450 ±20	700 ±30	370 ±20	94 ± 8	12 ±4	6 ±3	4 ±2	0.3 ±0.5
LLN 'omega'	3 ± 1	32 ± 4	71 ± 5	83 ± 8	45 ± 3	8 ±1	3.3 ±0.7	3 ±2	0.2 ±0.2
LLS 'zeta'	58 ± 6	490 ±20	850 ±30	440 ±30	110 ±10	13 ±4	5 ±2	5 ±1	0.6 ±0.7
LLS 'omega'	3 ± 1	39 ± 4	94 ± 5	97 ± 8	51 ± 5	10 ±2	5 ±0.8	4 ±2	0.7 ±0.4

Table C.5 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, P2. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	$10^{-0.4}$-$10^{-0.2}$ $R_\oplus$	$10^{-0.2}$-10^{0} $R_\oplus$	10^{0}-$10^{.2}$ $R_\oplus$	$10^{.2}$-$10^{.4}$ $R_\oplus$	$10^{.4}$-$10^{.6}$ $R_\oplus$	$10^{.6}$-$10^{.8}$ $R_\oplus$	$10^{.8}$-10^{1} $R_\oplus$	10^{1}-$10^{1.2}$ $R_\oplus$	$10^{1.2}$-$10^{1.4}$ $R_\oplus$
Observed *Kepler* DR25	0	0	0	0	0	0	0	0	0
P2									
32 unblended									
LLN 'zeta'	10 ±3	27 ±6	21 ±4	6 ±2	1 ±2	0.1 ±0.3	0	0.3 ±0.5	0
LLN 'omega'	0.8 ±0.5	1.3 ±0.9	1.7 ±0.4	1.8 ±0.8	0.8 ±0.8	0.2 ±0.2	0	0.1 ±0.2	0
LLS 'zeta'	11 ±4	31 ±6	23 ±4	9 ±2	2 ±2	0.1 ±0.3	0.2 ±0.4	0	0
LLS 'omega'	0.5 ±0.6	1.9 ±0.4	2 ±1	1.0 ±0.7	0.6 ±1.0	0.1 ±0.2	0	0.1 ±0.2	0
32 inc blends									
LLN 'zeta'	10 ±3	27 ±6	20 ±4	6 ±2	1 ±1	0.1 ±0.3	0	0.3 ±0.5	0
LLN 'omega'	0.8 ±0.5	1.3 ±0.9	1.7 ±0.4	1.9 ±0.9	0.7 ±0.6	0.2 ±0.2	0	0.1 ±0.2	0
LLS 'zeta'	11 ±4	32 ±7	22 ±5	9 ±2	2 ±2	0.1 ±0.3	0.2 ±0.4	0	0
LLS 'omega'	0.5 ±0.6	2.0 ±0.5	1.9 ±1.2	1.0 ±0.7	0.6 ±1.0	0.1 ±0.2	0 ±0	0.1 ±0.2	0
28 unblended									
LLN 'zeta'	9 ±3	26 ±6	20 ±4	6 ±2	1 ±1	0.1 ±0.3	0	0.3 ±0.5	0
LLN 'omega'	0.8 ±0.5	1.2 ±0.8	1.7 ±0.4	1.8 ±0.8	0.8 ±0.8	0.2 ±0.2	0	0.1 ±0.2	0
LLS 'zeta'	10 ±4	30 ±6	23 ±4	9 ±2	2 ±2	0.1 ±0.3	0.2 ±0.4	0	0
LLS 'omega'	0.4 ±0.4	1.9 ±0.4	2 ±1	1.0 ±0.7	0.6 ±1.0	0.1 ±0.2	0	0.1 ±0.2	0
28 inc blends									
LLN 'zeta'	9 ±2	26 ±6	20 ±4	6 ±2	1 ±1	0.1 ±0.3	0	0.3 ±0.5	0
LLN 'omega'	0.8 ±0.5	1.2 ±0.8	1.7 ±0.4	1.9 ±0.9	0.7 ±0.6	0.2 ±0.2	0	0.1 ±0.2	0
LLS 'zeta'	10 ±4	30 ±7	22 ±4	9 ±2	2 ±2	0.1 ±0.3	0.2 ±0.4	0	0
LLS 'omega'	0.4 ±0.4	2.0 ±0.5	1.9 ±1.2	1.0 ±0.7	0.6 ±1.0	0.1 ±0.2	0	0.1 ±0.2	0
24 unblended									
LLN 'zeta'	8 ±2	25 ±6	20 ±4	6 ±2	1 ±1	0.1 ±0.3	0	0.3 ±0.5	0
LLN 'omega'	0.6 ±0.4	1.1 ±0.9	1.6 ±0.4	1.8 ±0.8	0.8 ±0.8	0.2 ±0.2	0	0.1 ±0.2	0
LLS 'zeta'	9 ±4	29 ±7	23 ±4	9 ±2	2 ±2	0.1 ±0.3	0.2 ±0.3	0	0
LLS 'omega'	0.3 ±0.3	1.9 ±0.4	2 ±1	1.0 ±0.7	0.6 ±1.0	0.1 ±0.2	0	0.1 ±0.2	0
24 inc blends									
LLN 'zeta'	8 ±2	24 ±7	20 ±4	6 ±2	1 ±1	0.1 ±0.3	0	0.3 ±0.5	0
LLN 'omega'	0.6 ±0.4	1.1 ±0.9	1.6 ±0.4	1.9 ±0.9	0.7 ±0.6	0.2 ±0.2	0	0.1 ±0.2	0
LLS 'zeta'	9 ±4	29 ±7	22 ±4	9 ±2	2 ±2	0.1 ±0.3	0.2 ±0.4	0	0
LLS 'omega'	0.3 ±0.3	2.0 ±0.5	1.9 ±1.2	1.0 ±0.7	0.6 ±1.0	0.1 ±0.2	0	0.1 ±0.2	0

Table C.6 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, P4. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	$10^{-0.4}$-$10^{-0.2}$ $R_\oplus$	$10^{-0.2}$-10^{0} $R_\oplus$	10^{0}-$10^{.2}$ $R_\oplus$	$10^{.2}$-$10^{.4}$ $R_\oplus$	$10^{.4}$-$10^{.6}$ $R_\oplus$	$10^{.6}$-$10^{.8}$ $R_\oplus$	$10^{.8}$-10^{1} $R_\oplus$	10^{1}-$10^{1.2}$ $R_\oplus$	$10^{1.2}$-$10^{1.4}$ $R_\oplus$
Observed *Kepler* DR25	1	14	38	26	3	1	1	1	0
P4									
32 unblended									
LLN 'zeta'	35 ±6	240 ±10	480 ±20	280 ±20	68 ± 9	8 ±3	3 ±2	4 ±1	0.1 ±0.3
LLN 'omega'	1.2 ±0.9	18 ± 3	51 ± 8	66 ± 6	32 ± 5	6 ±1	2 ±1	3 ±1	0.7 ±0.5
LLS 'zeta'	44 ±7	270 ±10	550 ±20	320 ±20	75 ± 7	10 ±3	4 ±2	4 ±1	0.2 ±0.4
LLS 'omega'	3 ±2	17.6 ±0.7	57 ± 5	71 ± 2	39 ± 4	8 ±2	2.8 ±0.8	2.2 ±0.6	0.3 ±0.4
32 inc blends									
LLN 'zeta'	35 ±6	220 ±10	410 ±10	240 ±10	60 ±10	8 ±2	3 ±2	3 ±2	0.1 ±0.3
LLN 'omega'	1 ±1	20 ± 2	47 ± 7	57 ± 5	26 ± 4	5 ±1	1.9 ±0.6	2.7 ±0.7	0.4 ±0.4
LLS 'zeta'	43 ±8	260 ±20	480 ±20	280 ±20	65 ± 5	9 ±3	3 ±2	4 ±1	0.2 ±0.4
LLS 'omega'	3 ±2	17 ± 1	53 ± 4	67 ± 3	33 ± 3	7 ±2	3 ±1	2.0 ±0.6	0.3 ±0.4
28 unblended									
LLN 'zeta'	30 ±5	210 ±10	450 ±10	280 ±20	69 ± 9	8 ±3	3 ±2	4 ±2	0.1 ±0.3
LLN 'omega'	1.0 ±0.8	17 ± 3	47 ± 8	64 ±6	32 ± 5	6 ±1	2 ±1	3 ±1	0.7 ±0.5
LLS 'zeta'	38 ±6	240 ±10	510 ±20	310 ±20	75 ± 7	10 ±3	4 ±2	4 ±2	0.2 ±0.4
LLS 'omega'	3 ±2	15.5 ±0.3	53 ± 4	70 ± 2	39 ± 4	8 ±2	2.8 ±0.8	2.2 ±0.6	0.3 ±0.4
28 inc blends									
LLN 'zeta'	29 ±5	200 ±10	380 ±10	240 ±20	55 ± 9	8 ±2	3 ±2	3 ±2	0.1 ±0.3
LLN 'omega'	1 ±1	18 ± 2	43 ± 7	56 ± 5	26 ± 4	5 ±1	1.9 ±0.6	2.7 ±0.9	0.4 ±0.4
LLS 'zeta'	37 ±7	230 ±20	440 ±20	280 ±20	65 ± 5	9 ±3	3 ±2	4 ±1	0.2 ±0.4
LLS 'omega'	3 ±2	15 ± 1	50 ± 3	66 ± 2	33 ± 3	7 ±2	3 ±1	2 ±0.6	0.3 ±0.4
24 unblended									
LLN 'zeta'	25 ±4	180 ±10	410 ±10	270 ±20	69 ± 9	8 ±3	3 ±2	4 ±2	0.1 ±0.3
LLN 'omega'	0.8 ±0.8	15 ± 3	43 ± 8	63 ± 5	32 ± 5	6 ±1	2 ±1	2.8 ±1	0.7 ±0.5
LLS 'zeta'	32 ±5	207 ±10	460 ±20	300 ±20	75 ± 7	10 ±3	4 ±2	4 ±1	0.2 ±0.4
LLS 'omega'	2 ±2	13 ±0.5	48 ± 5	68 ± 2	39 ±4	7 ±2	2.8 ±0.8	2.2 ±0.6	0.3 ±0.4
24 inc blends									
LLN 'zeta'	23 ±4	170 ±10	350 ±10	230 ±10	55 ± 9	8 ±2	3 ±2	3 ±2	0.1 ±0.3
LLN 'omega'	0.8 ±0.9	15 ± 3	40 ± 8	55 ± 5	26 ± 4	5 ±1	1.9 ±0.6	2.7 ±0.9	0.4 ±0.4
LLS 'zeta'	31 ±7	200 ±20	410 ±10	270 ±20	65 ± 5	9 ±3	3 ±2	4 ±1	0.2 ±0.4
LLS 'omega'	2 ±2	13 ± 1	46 ± 4	64 ± 3	33 ± 3	7 ±2	3 ±1	2.0 ±0.6	0.3 ±0.4

Table C.7 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, P5. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	$10^{-0.4}$-$10^{-0.2}$ $R_\oplus$	$10^{-0.2}$-10^{0} $R_\oplus$	10^{0}-$10^{.2}$ $R_\oplus$	$10^{.2}$-$10^{.4}$ $R_\oplus$	$10^{.4}$-$10^{.6}$ $R_\oplus$	$10^{.6}$-$10^{.8}$ $R_\oplus$	$10^{.8}$-10^{1} $R_\oplus$	10^{1}-$10^{1.2}$ $R_\oplus$	$10^{1.2}$-$10^{1.4}$ $R_\oplus$
Observed *Kepler* DR25	6	44	84	74	55	14	10	6	4
P5									
32 unblended									
LLN 'zeta'	102 ± 9	1180 ±40	3290 ±90	2740 ±50	900 ±40	120 ±10	47 ±7	40 ±5	4 ±2
LLN 'omega'	5 ± 2	81 ± 2	360 ±10	610 ±20	450 ±10	87 ± 5	37 ±3	33 ±3	4.5 ±0.9
LLS 'zeta'	110 ±10	1310 ±30	3810 ±50	3130 ±70	990 ±20	130 ±10	49 ±6	46 ±6	4 ±2
LLS 'omega'	7 ± 2	94 ± 8	414 ± 4	700 ±10	480 ±20	105 ± 2	38 ±2	40 ±3	7 ±2
32 inc blends									
LLN 'zeta'	102 ± 9	1220 ±50	3140 ±70	2690 ±50	840 ±40	110 ±10	48 ±7	38 ±5	4 ±2
LLN 'omega'	5 ± 2	88 ± 3	350 ±10	620 ±30	420 ±10	85 ± 5	37 ±3	31 ±3	4.2 ±0.5
LLS 'zeta'	109 ± 9	1360 ±30	3640 ±50	3100 ±70	930 ±20	130 ±10	51 ±5	42 ±6	4 ±2
LLS 'omega'	7 ± 2	101 ± 9	401 ± 6	700 ±10	450 ±20	103 ± 2	40 ±1	37 ±4	7 ±1
28 unblended									
LLN 'zeta'	83 ± 7	1010 ±30	3000 ±80	2610 ±50	890 ±40	120 ±10	47 ±7	40 ±5	4 ±2
LLN 'omega'	4 ± 2	71 ± 3	330 ±10	590 ±20	440 ±10	86 ± 5	37 ±3	33 ±3	4.5 ±0.9
LLS 'zeta'	89 ± 7	1120 ±30	3460 ±50	3000 ±60	980 ±20	130 ±10	49 ±6	46 ±6	4 ±2
LLS 'omega'	6 ± 2	79 ± 7	379 ± 4	670 ±10	470 ±20	105 ± 2	38 ±2	40 ±3	7 ±2
28 inc blends									
LLN 'zeta'	82 ± 7	1050 ±30	2860 ±70	2570 ±50	820 ±40	110 ±10	48 ±7	37 ±5	4 ±2
LLN 'omega'	4 ± 2	76 ± 4	320 ±10	591 ±20	410 ±10	85 ± 5	37 ±3	31 ±3	4.2 ±0.5
LLS 'zeta'	88 ± 7	1160 ±50	3320 ±50	2960 ±60	920 ±20	130 ±10	50 ±5	42 ±6	4 ±2
LLS 'omega'	6 ± 2	84 ± 7	367 ± 7	670 ±10	440 ±20	103 ± 2	40 ±1	37 ±4	7 ±1
24 unblended									
LLN 'zeta'	64 ± 8	840 ±30	2680 ±80	2470 ±50	870 ±40	110 ±10	47 ±7	40 ±5	4 ±2
LLN 'omega'	3 ± 1	59 ± 2	300 ±10	550 ±20	430 ±10	85 ± 5	37 ±3	33 ±3	4.5 ±0.9
LLS 'zeta'	72 ± 8	930 ±30	3090 ±40	2830 ±60	960 ±20	130 ±10	49 ±6	46 ±6	4 ±2
LLS 'omega'	4 ± 1	67 ± 6	337 ± 6	626 ±10	470 ±20	105 ± 3	38 ±2	40 ±3	7 ±2
24 inc blends									
LLN 'zeta'	63 ± 8	880 ±30	2550 ±60	2430 ±40	800 ±40	110 ±10	48 ±7	37 ±5	4 ±2
LLN 'omega'	3 ± 1	64 ± 3	284 ± 9	550 ±20	400 ±10	84 ± 4	37 ±3	31 ±3	4.2 ±0.5
LLS 'zeta'	71 ± 8	970 ±30	2960 ±50	2800 ±60	900 ±20	130 ±10	51 ±5	42 ±6	4 ±2
LLS 'omega'	4 ± 1	72 ±6	330 ±10	630 ±10	440 ±20	102 ± 3	40 ±1	37 ±4	7 ±1

Table C.8 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, bins from Fressin et al. (2013), $V < 26$. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	0.4-0.8 $R_\oplus$	0.8-1.25 $R_\oplus$	1.25-2 $R_\oplus$	2-4 $R_\oplus$	4-6 $R_\oplus$	6-22 $R_\oplus$	22-25 $R_\oplus$
Observed *Kepler* DR25	51	245	698	1017	120	152	2
$V < 26$							
32 unblended							
LLN 'zeta'	610 ±30	2890 ±50	7820 ±90	12700 ±100	1640 ±30	4280 ±50	140 ±20
LLN 'omega'	40 ± 4	270 ±10	1260 ±10	4640 ± 70	1230 ±10	3620 ±30	153 ± 7
LLS 'zeta'	670 ±20	3320 ±60	8760 ±60	13200 ±100	1520 ±50	3170 ±40	70 ± 8
LLS 'omega'	46 ± 6	300 ±10	1410 ±20	4820 ± 60	1150 ±30	2620 ±20	80 ± 7
32 inc blends							
LLN 'zeta'	600 ±30	2840 ±40	7290 ±90	10750 ±100	1380 ±30	2130 ±40	16 ± 5
LLN 'omega'	41 ± 5	270 ±10	1200 ± 6	3960 ± 70	1040 ±20	1800 ± 8	15 ± 2
LLS 'zeta'	660 ±30	3260 ±60	8270 ±60	11610 ± 90	1340 ±40	1940 ±30	16 ± 4
LLS 'omega'	46 ± 6	300 ±10	1360 ±30	4260 ± 50	1010 ±30	1610 ±30	16 ± 3
28 unblended							
LLN 'zeta'	500 ±30	2510 ±40	6970 ±80	11700 ±200	1550 ±30	4110 ±50	130 ±20
LLN 'omega'	33 ± 5	240 ±10	1130 ±10	4280 ± 70	1160 ±10	3470 ±30	147 ± 6
LLS 'zeta'	560 ±20	2850 ±50	7830 ±60	12180 ±100	1460 ±50	3070 ±30	67 ± 9
LLS 'omega'	38 ± 6	260 ±10	1260 ±20	4460 ± 50	1090 ±30	2550 ±30	77 ± 7
28 inc blends							
LLN 'zeta'	500 ±30	2470 ±40	6500 ±70	9900 ±100	1320 ±20	2090 ±40	16 ±5
LLN 'omega'	34 ± 5	230 ±10	1076 ± 7	3670 ± 60	1000 ±20	1770 ± 9	15 ±2
LLS 'zeta'	550 ±20	2810 ±50	7410 ±60	10770 ± 70	1290 ±40	1910 ±30	15 ±4
LLS 'omega'	38 ± 6	260 ±10	1220 ±30	3960 ± 50	970 ±30	1580 ±30	16 ±3
24 unblended							
LLN 'zeta'	410 ±20	2130 ±40	6080 ±70	10600 ±100	1460 ±20	3920 ±50	120 ±20
LLN 'omega'	28 ± 5	200 ±10	990 ±10	3890 ± 50	1090 ±20	3300 ±30	141 ± 5
LLS 'zeta'	450 ±20	2410 ±50	6860 ±60	11100 ±100	1380 ±50	2970 ±30	65 ± 8
LLS 'omega'	31 ± 5	230 ± 10	1100 ±20	4060 ± 50	1040 ±20	2430 ±50	75 ± 7
24 inc blends							
LLN 'zeta'	400 ±20	2090 ±40	5690 ±50	9100 ±100	1250 ±20	2040 ±40	16 ± 5
LLN 'omega'	29 ± 5	200 ±10	943 ± 4	3350 ± 50	940 ±20	1728 ± 9	15 ± 2
LLS 'zeta'	440 ±20	2370 ±50	6510 ±70	9850 ± 80	1230 ±40	1880 ±30	15 ± 4
LLS 'omega'	31 ± 5	220 ±10	1070 ±30	3630 ± 50	930 ±20	1560 ±30	16 ± 3

Table C.9 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, bins from Fressin et al. (2013), $V < 16$, F5-M. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	0.4-0.8 $R_\oplus$	0.8-1.25 $R_\oplus$	1.25-2 $R_\oplus$	2-4 $R_\oplus$	4-6 $R_o plus$	6-22 $R_\oplus$	22-25 $R_\oplus$
Observed *Kepler* DR25	49	244	692	1006	117	146	1
$V < 16$, F5-M							
32 unblended							
LLN 'zeta'	560 ±30	2630 ±50	6690 ±90	10090 ±100	1070 ±20	1570 ±30	20 ±5
LLN 'omega'	36 ± 4	250 ±10	1070 ±10	3630 ± 70	810 ±10	1300 ±20	21 ±2
LLS 'zeta'	630 ±20	3010 ±70	7570 ±60	10670 ±100	1050 ±40	1440 ±30	15 ±3
LLS 'omega'	44 ± 5	275 ± 9	1200 ±20	3830 ± 30	800 ±20	1170 ±30	15 ±3
32 inc blends							
LLN 'zeta'	560 ±30	2580 ±50	6260 ±80	8700 ±100	980 ±20	1270 ±30	11 ±4
LLN 'omega'	38 ± 5	245 ± 9	1015 ± 4	3180 ± 70	740 ±10	1060 ±20	10.2 ±0.6
LLS 'zeta'	620 ±30	2960 ±60	7170 ±60	9560 ± 80	980 ±40	1240 ±30	11 ±3
LLS 'omega'	44 ± 5	270 ±10	1160 ±20	3450 ± 30	750 ±20	1020 ±20	11 ±2
28 unblended							
LLN 'zeta'	470 ±30	2290 ±40	5990 ±80	9300 ±100	1030 ±20	1560 ±30	20 ±5
LLN 'omega'	31 ± 5	210 ±10	960 ±20	3370 ± 70	780 ±10	1290 ±20	21 ±2
LLS 'zeta'	530 ±20	2600 ±60	6790 ±60	9900 ± 90	1010 ±40	1430 ±30	15 ±3
LLS 'omega'	37 ± 6	239 ± 9	1080 ±20	3560 ± 30	770 ±30	1170 ±30	15 ±3
28 inc blends							
LLN 'zeta'	470 ±30	2250 ±40	5610 ±60	8100 ±100	950 ±20	1260 ±30	11 ±4
LLN 'omega'	31 ± 4	210 ±10	917 ± 7	2960 ±60	720 ±20	1050 ±20	10.2 ±0.6
LLS 'zeta'	520 ±20	2560 ±50	6440 ±60	8900 ± 70	950 ±40	1230 ±30	11 ±3
LLS 'omega'	36 ± 5	240 ±10	1050 ±20	3220 ± 30	730 ±20	1010 ±20	11 ±2
24 unblended							
LLN 'zeta'	380 ±20	1950 ±40	5260 ±60	8500 ±100	990 ±20	1540 ±30	20 ±5
LLN 'omega'	26 ± 5	180 ±10	840 ±10	3090 ± 60	750 ±10	1270 ±20	21 ±2
LLS 'zeta'	430 ±20	2200 ±50	5970 ±70	9060 ± 80	970 ±30	1420 ±30	15 ±3
LLS 'omega'	29 ± 5	204 ± 8	950 ±20	3270 ± 30	740 ±20	1150 ±30	15 ±3
24 inc blends							
LLN 'zeta'	370 ±20	1910 ±40	4940 ±40	7400 ±100	910 ±20	1250 ±30	11 ±4
LLN 'omega'	27 ± 5	180 ±10	807 ± 6	2710 ± 60	690 ±10	1040 ±20	10.2 ±0.6
LLS 'zeta'	420 ±20	2170 ±50	5680 ±70	8170 ± 70	920 ±40	1220 ±30	11 ±3
LLS 'omega'	30 ± 5	203 ± 8	930 ±20	2960 ± 30	700 ±20	1000 ±20	11 ±2

Table C.10 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, bins from Fressin et al. (2013), P1. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	0.4-0.8 $R_\oplus$	0.8-1.25 $R_\oplus$	1.25-2 $R_\oplus$	2-4 $R_\oplus$	4-6 $R_\oplus$	6-22 $R_\oplus$	22-25 $R_\oplus$
Observed *Kepler* DR25	6	3	10	7	2	1	0
P1							
32 unblended							
LLN 'zeta'	270 ±20	690 ±20	680 ±30	290 ±20	10 ±4	12 ±3	0
LLN 'omega'	18 ± 4	59 ± 5	94 ± 5	88 ± 6	7 ±2	8 ±2	0
LLS 'zeta'	300 ±10	790 ±30	810 ±40	340 ±30	11 ±4	11 ±2	0.2 ±0.4
LLS 'omega'	19 ± 4	74 ± 5	114 ± 5	104 ± 3	9 ±2	11 ±1	0.3 ±0.4
32 inc blends							
LLN 'zeta'	270 ±20	690 ±40	680 ±30	290 ±10	10 ±4	12 ±3	0
LLN 'omega'	18 ± 4	58 ± 4	94 ± 5	87 ± 5	7 ±1	8 ±2	0
LLS 'zeta'	290 ±20	790 ±30	800 ±30	340 ±30	12 ±4	11 ±2	0.2 ±0.4
LLS 'omega'	20 ± 4	74 ± 5	114 ± 5	103 ± 3	9 ±1	11 ±1	0.2 ±0.3
28 unblended							
LLN 'zeta'	230 ±20	630 ±30	660 ±30	290 ±10	10 ±4	12 ±3	0
LLN 'omega'	16 ± 4	53 ± 5	90 ± 4	87 ± 5	7 ±2	8 ±1	0
LLS 'zeta'	250 ±10	720 ±30	790 ±30	340 ±30	11 ±4	11 ±2	0.2 ±0.4
LLS 'omega'	17 ± 4	68 ± 5	112 ± 5	102 ± 3	9 ±2	11 ±1	0.2 ±0.3
28 inc blends							
LLN 'zeta'	230 ±20	630 ±20	650 ±30	280 ±10	10 ±4	12 ±3	0
LLN 'omega'	16 ± 4	53 ± 5	90 ± 5	86 ± 5	7 ±1	8 ±2	0
LLS 'zeta'	250 ±10	720 ±30	780 ±30	340 ±30	12 ±4	11 ±2	0.2 ±0.4
LLS 'omega'	18 ± 4	68 ± 5	111 ± 6	102 ± 3	9 ±1	11 ±1	0.2 ±0.3
24 unblended							
LLN 'zeta'	190 ±20	570 ±30	630 ±20	290 ±10	10 ±4	12 ±3	0
LLS 'omega'	14 ± 4	48 ± 4	87 ± 5	86 ± 6	7 ±2	8 ±1	0
LLS 'zeta'	210 ±10	650 ±30	750 ±30	340 ±30	11 ±4	11 ±2	0.2 ±0.4
LLS 'omega'	14 ± 4	63 ± 5	107 ± 6	101 ± 2	9 ±2	11 ± 1	0.2 ±0.3
24 inc blends							
LLN 'zeta'	190 ±20	570 ±30	630 ±20	280 ±10	10 ±4	12 ±3	0
LLN 'omega'	14 ± 4	48 ± 4	87 ± 6	85 ± 5	7 ±1	8 ±2	0
LLS 'zeta'	210 ±10	650 ±30	750 ±30	330 ±30	12 ±4	11 ±2	0.2 ±0.4
LLS 'omega'	14 ± 4	62 ± 5	107 ± 6	100 ± 3	9 ±1	11 ±1	0.2 ±0.3

Table C.11 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, bins from Fressin et al. (2013), P2. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	0.4-0.8 $R_\oplus$	0.8-1.25 $R_\oplus$	1.25-2 $R_\oplus$	2-4 $R_\oplus$	4-6 $R_\oplus$	6-22 $R_\oplus$	22-25 $R_\oplus$
Observed *Kepler* DR25	0	0	0	0	0	0	0
P2							
32 unblended							
LLN 'zeta'	22 ±4	24 ±7	15 ±4	4 ±2	0.1 ±0.3	0.3 ±0.5	0
LLN 'omega'	1 ±1	1.5 ±0.7	2.0 ±0.6	2 ±1	0.2 ±0.2	0.1 ±0.2	0
LLS 'zeta'	24 ±6	27 ±5	18 ±2	6 ±3	0.1 ±0.3	0.2 ±0.4	0
LLS 'omega'	1.5 ±0.7	2 ±1	1.3 ±0.3	1 ±1	0.1 ±0.2	0.1 ±0.2	0
32 inc blends							
LLN 'zeta'	22 ±4	24 ±7	15 ±4	4 ±2	0.1 ±0.3	0.3 ±0.5	0
LLN 'omega'	1 ±1	1.5 ±0.7	2.0 ±0.6	2 ±1	0.2 ±0.2	0.1 ±0.2	0
LLS 'zeta'	24 ±6	27 ±5	18 ±3	6 ±3	0.1 ±0.3	0.2 ±0.4	0
LLS 'omega'	1.5 ±0.7	2 ±1	1.6 ±0.7	1.0 ±0.7	0.1 ±0.2	0.1 ±0.2	0
28 unblended							
LLN 'zeta'	21 ±3	23 ±7	15 ±4	4 ±2	0.1 ±0.3	0.3 ±0.5	0
LLN 'omega'	1 ±1	1.4 ±0.7	2.0 ±0.6	2 ±1	0.2 ±0.2	0.1 ±0.2	0
LLS 'zeta'	22 ±6	26 ±4	18 ±3	6 ±3	0.1 ±0.3	0.2 ±0.4	0
LLS 'omega'	1.4 ±0.6	2 ±1	1.3 ±0.3	1 ±1	0.1 ±0.2	0.1 ±0.2	0
28 inc blends							
LLN 'zeta'	20 ±3	23 ±7	14 ±4	4 ±2	0.10 ±0.3	0.3 ±0.5	0
LLN 'omega'	1 ±1	1.4 ±0.7	2.0 ±0.6	2 ±1	0.2 ±0.2	0.1 ±0.2	0
LLS 'zeta'	22 ±6	26 ±5	18 ±3	6 ±3	0.1 ±0.3	0.2 ±0.4	0
LLS 'omega'	1.4 ±0.6	2 ±1	1.3 ±0.3	1 ±1	0.1 ±0.2	0.1 ±0.2	0
24 unblended							
LLN 'zeta'	19 ±3	22 ±7	15 ±3	4 ±2	0.1 ±0.3	0.3 ±0.5	0
LLN 'omega'	1 ±1	1.3 ±0.8	2.0 ±0.6	2 ±1	0.2 ±0.2	0.1 ±0.2	0
LLS 'zeta'	21 ±5	25 ±5	18 ±3	6 ±3	0.1 ±0.3	0.2 ±0.4	0
LLS 'omega'	1.3 ±0.6	2 ±1	1.3 ±0.3	1 ±1	0.1 ±0.2	0.1 ±0.2	0
24 inc blends							
LLN 'zeta'	19 ±4	22 ±7	15 ±4	4 ±2	0.1 ±0.3	0.3 ±0.5	0
LLN 'omega'	1 ±1	1.3 ±0.8	2.0 ±0.6	2 ±1	0.2 ±0.2	0.1 ±0.2	0
LLS 'zeta'	21 ±5	25 ±5	18 ±3	6 ±3	0.1 ±0.3	0.2 ±0.4	0
LLS 'omega'	1.3 ±0.5	2 ±1	1.3 ±0.3	1 ±1	0.1 ±0.2	0.1 ±0.2	0

Table C.12 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, bins from Fressin et al. (2013), P4. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	0.4-0.8 $R_\oplus$	0.8-1.25 $R_\oplus$	1.25-2 $R_\oplus$	2-4 $R_\oplus$	4-6 $R_\oplus$	6-22 $R_\oplus$	22-25 $R_\oplus$
Observed *Kepler* DR25	6	28	36	12	1	2	0
P4							
32 unblended							
LLN 'zeta'	112 ± 9	330 ±20	450 ±20	210 ±10	7 ±3	8 ±3	0
LLN 'omega'	7 ± 2	31 ± 3	67 ±6	63 ± 6	5 ±1	6.2 ±0.4	0.3 ±0.3
LLS 'zeta'	130 ±10	380 ±20	510 ±20	230 ±20	8 ±3	9 ±3	0.1 ±0.3
LLS 'omega'	9 ± 3	31 ± 3	74 ± 6	73 ± 4	7 ±2	6 ±1	0.1 ±0.2
32 inc blends							
LLN 'zeta'	110 ± 6	310 ±20	380 ±20	170 ±10	7 ±2	7 ±3	0
LLN 'omega'	8 ± 2	30 ± 2	60 ± 4	53 ± 5	5.0 ±0.8	5.1 ±0.7	0.2 ±0.3
LLS 'zeta'	130 ±10	350 ±20	450 ±20	200 ±10	8 ±3	8 ±3	0
LLS 'omega'	9 ± 3	30 ± 3	70 ± 5	60 ± 4	6 ±1	5 ±1	0.1 ±0.2
28 unblended							
LLN 'zeta'	97 ± 7	300 ±20	430 ±20	210 ±10	7 ±3	8 ±3	0
LLN 'omega'	6 ± 2	30 ± 3	63 ± 6	63 ± 6	5 ±1	6.2 ±0.4	0.3 ±0.3
LLS 'zeta'	115 ± 9	340 ±20	490 ±20	230 ±20	8 ±3	9 ±3	0.1 ±0.3
LLS 'omega'	8 ± 2	28 ± 3	72 ± 6	73 ± 4	7 ±2	6 ±1	0.1 ±0.2
28 inc blends							
LLN 'zeta'	95 ± 5	280 ±20	360 ±20	160 ±10	7 ±2	7 ±3	0
LLN 'omega'	7 ± 2	27 ± 3	57 ± 4	53 ± 5	5.0 ±0.8	5.1 ±0.7	0.2 ±0.3
LLS 'zeta'	110 ±10	320 ±20	430 ±20	200 ±10	8 ±3	8 ±3	0
LLS 'omega'	8 ± 2	27 ± 3	68 ± 5	64 ± 4	6 ±1	5 ±1	0.1 ±0.2
24 unblended							
LLN 'zeta'	83 ± 7	270 ±10	400 ±20	200 ±10	7 ±3	8 ±3	0
LLN 'omega'	5 ± 1	26 ± 4	60 ± 6	62 ± 6	5 ±1	6.2 ±0.4	0.3 ±0.3
LLS 'zeta'	99 ± 9	300 ±20	450 ±20	230 ±20	8 ±3	9 ±3	0.1 ±0.3
LLS 'omega'	7 ± 2	24 ± 3	68 ± 6	72 ±5	7 ±2	6 ±1	0.1 ±0.2
24 inc blends							
LLN 'zeta'	81 ± 7	250 ±20	337 ±20	160 ±10	7 ±2	7 ±3	0
LLN 'omega'	6 ± 2	24 ± 4	54 ± 5	52 ± 5	5.0 ±0.8	5.1 ±0.7	0.2 ±0.3
LLS 'zeta'	97 ± 9	280 ±20	400 ±10	200 ±10	8 ±3	8 ±3	0
LLS 'omega'	7 ± 2	23 ± 3	64 ± 5	64 ± 4	6 ±1	5 ±1	0.1 ±0.2

Table C.13 Numbers of observable planets from simulated PLATO Long Look fields, in the *log R log P* plane, bins from Fressin et al. (2013), P5. Red indicates highest value in a given row. Error is σ.

Population cameras With blends?	0.4-0.8 $R_\oplus$	0.8-1.25 $R_\oplus$	1.25-2 $R_\oplus$	2-4 $R_\oplus$	4-6 $R_\oplus$	6-22 $R_\oplus$	22-25 $R_\oplus$
Observed *Kepler* DR25	23	50	91	99	12	22	0
P5							
32 unblended							
LLN 'zeta'	440 ±30	1910 ±40	3450 ±70	2410 ±70	100 ±10	102 ± 8	0.6 ±0.8
LLN 'omega'	27 ± 4	174 ± 6	524 ± 6	780 ±20	76 ± 6	84 ± 4	0.7 ±0.4
LLS 'zeta'	480 ±20	2180 ±50	3990 ±80	2700 ±60	110 ±10	115 ± 7	0.9 ±0.7
LLS 'omega'	33 ± 5	202 ± 8	590 ±10	870 ±20	91 ± 2	98 ± 5	1.5 ±0.8
32 inc blends							
LLN 'zeta'	430 ±30	1890 ±50	3340 ±70	2320 ±60	100 ±10	98 ± 8	0.5 ±0.8
LLN 'omega'	28 ± 4	174 ± 5	514 ±6	760 ±10	77 ± 6	80 ± 4	0.7 ±0.4
LLS 'zeta'	480 ±20	2170 ±50	3870 ±70	2620 ±60	110 ±10	111 ± 8	0.9±0.7
LLS 'omega'	34 ± 5	200 ± 9	580 ±10	840 ±20	91 ± 2	95 ± 5	1.5 ±0.8
28 unblended							
LLN 'zeta'	360 ±20	1690 ±40	3190 ±60	2340 ±70	100 ±10	101 ± 7	0.6 ±0.8
LLN 'omega'	24 ± 4	154 ± 5	489 ± 8	760 ±20	75 ± 5	84 ± 4	0.7 ±0.4
LLS 'zeta'	400 ±20	1910 ±50	3700 ±70	2640 ±60	110 ±10	115 ± 7	0.9 ±0.7
LLS 'omega'	28 ± 5	180 ± 8	550 ±30	840 ±20	91 ± 2	98 ± 5	1.5 ±0.8
28 inc blends							
LLN 'zeta'	360 ±20	1670 ±30	3100 ±60	2250 ±60	100 ±10	98 ± 8	0.5 ±0.8
LLN 'omega'	23 ± 3	154 ± 6	480 ± 8	740 ±10	76 ± 6	80 ± 4	0.7 ±0.4
LLS 'zeta'	400 ±20	1900 ±50	3600 ±70	2550 ±50	110 ±10	111 ± 8	0.9 ±0.7
LLS 'omega'	28 ± 5	177 ± 9	540 ±20	830 ±20	91 ± 2	95 ± 5	1.5 ±0.8
24 unblended							
LLN 'zeta'	290 ±20	1440 ±40	2920 ±60	2260 ±70	100 ±10	102 ± 8	0.6 ±0.8
LLN 'omega'	20 ± 4	132 ± 4	449 ± 8	740 ±20	75 ± 5	84 ± 4	0.7 ±0.4
LLS 'zeta'	320 ±20	1640 ±40	3370 ±70	2550 ±50	112 ±10	115 ± 7	0.9 ±0.7
LLS 'omega'	22 ± 4	156 ± 6	500 ±10	820 ±10	91 ± 2	98 ± 5	1.5 ±0.8
24 inc blends							
LLN 'zeta'	290 ±20	1440 ±40	2830 ±50	2170 ±60	100 ±10	98 ± 8	0.5 ±0.8
LLN 'omega'	20 ± 4	130 ± 5	439 ± 7	720 ±10	76 ± 6	80 ± 4	0.7 0.4
LLS 'zeta'	320 ±20	1630 ±40	3290 ±70	2460 ±50	112 ±10	111 ± 8	0.9 ±0.7
LLS 'omega'	22 ± 4	155 ± 7	500 ±20	800 ±10	90 ± 3	95 ± 5	1.5 ±0.8

Appendix D

Detailed results of binary simulations

Table D.1 Observable binary population: unblended, $V < 26$, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	26900 ±200	26800 ±200	26800 ±200
$[-0.6, -0.4]$	0	0	0
$[-0.4, -0.2]$	0	0	0
$[-0.2, 0.0]$	0.5 ±0.6	0.5 ±0.6	0.5 ±0.6
$[0.0, 0.2]$	25 ± 5	23 ± 5	20 ± 5
$[0.2, 0.4]$	146 ± 8	133 ± 7	117 ± 7
$[0.4, 0.6]$	185 ± 8	172 ± 8	159 ± 8
$[0.6, 0.8]$	420 ± 20	400 ± 20	390 ± 20
$[0.8, 1.0]$	780 ± 20	780 ± 10	770 ± 20
$[1.0, 1.2]$	1480 ± 60	1480 ± 60	1470 ± 60
$[1.2, 1.4]$	2460 ± 20	2460 ± 20	2460 ± 20
$[1.4, 1.6]$	3900 ± 60	3900 ± 60	3900 ± 60
$[1.6, 1.8]$	6020 ± 60	6020 ± 60	6020 ± 60
$[1.8, 2.0]$	7350 ± 70	7350 ± 70	7350 ± 70

Table D.2 Observable binary population: unblended, $V < 26$, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	22300 ±300	22200 ±300	22200±300
$[-0.6, -0.4]$	0	0	0
$[-0.4, -0.2]$	0	0	0
$[-0.2, 0.0]$	1 ± 1	0.7 ±0.7	0.7 ±0.7
$[0.0, 0.2]$	30 ± 4	27 ± 4	24 ± 3
$[0.2, 0.4]$	150 ± 10	135 ± 8	119 ± 6
$[0.4, 0.6]$	190 ± 3	182 ± 4	167 ± 4
$[0.6, 0.8]$	360 ± 20	350 ± 10	340 ± 10
$[0.8, 1.0]$	720 ± 10	717 ± 9	710 ± 9
$[1.0, 1.2]$	1280 ± 30	1280 ± 30	1280 ± 20
$[1.2, 1.4]$	2170 ± 30	2170 ± 30	2170 ± 30
$[1.4, 1.6]$	3500 ±100	3500 ±100	3500 ±100
$[1.6, 1.8]$	5200 ±100	5200 ±100	5200 ±100
$[1.8, 2.0]$	6000 ± 80	6000 ± 80	6000 ± 80

Table D.3 Observable binary population: same pixel, $V < 26$, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	44100 ±200	43500 ±200	42900 ±200
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	1 ± 1	1 ± 1	1 ± 1
$[-0.2,0.0]$	10 ± 4	8 ± 3	7 ± 2
$[0.0,0.2]$	110 ± 20	100 ± 20	80 ± 10
$[0.2,0.4]$	530 ± 10	470 ± 10	410 ± 10
$[0.4,0.6]$	1410 ± 30	1290 ± 30	1140 ± 30
$[0.6,0.8]$	3020 ± 20	2850 ± 20	2640 ± 20
$[0.8,1.0]$	4490 ± 70	4350 ± 60	4190 ± 70
$[1.0,1.2]$	4910 ± 80	4860 ± 80	4810 ± 70
$[1.2,1.4]$	4910 ± 50	4910 ± 50	4900 ± 50
$[1.4,1.6]$	5460 ± 60	5460 ± 60	5460 ± 60
$[1.6,1.8]$	6900 ± 60	6900 ± 60	6900 ± 60
$[1.8,2.0]$	7810 ± 70	7810 ± 70	7810 ± 70

Table D.4 Observable binary population: same pixel, $V < 26$, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	29700 ±300	29500 ±300	29200 ±300
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0.8 ±0.7	0.7 ±0.8	0.5 ±0.8
$[-0.2,0.0]$	8 ± 2	7 ± 1	7 ± 1
$[0.0,0.2]$	75 ± 7	65 ± 5	58 ± 5
$[0.2,0.4]$	360 ± 10	321 ± 7	285 ± 4
$[0.4,0.6]$	829 ± 7	763 ± 8	690 ± 6
$[0.6,0.8]$	1620 ± 50	1540 ± 40	1450 ± 40
$[0.8,1.0]$	2300 ± 20	2250 ± 20	2190 ± 30
$[1.0,1.2]$	2610 ± 10	2600 ± 10	2590 ± 20
$[1.2,1.4]$	3110 ± 40	3110 ± 40	3100 ± 30
$[1.4,1.6]$	4100 ±100	4100 ±100	4100 ±100
$[1.6,1.8]$	5600 ±100	5600 ±100	5600 ±100
$[1.8,2.0]$	6220 ± 80	6220 ± 80	6220 ± 80

Table D.5 Observable binary population: same imagette, $V < 26$, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	638800 ±500	620600 ±500	600100 ±400
$[-0.6, -0.4]$	2 ± 1	1 ± 1	0.9 ±0.5
$[-0.4, -0.2]$	35 ± 4	28 ± 3	22 ± 3
$[-0.2, 0.0]$	380 ± 20	320 ± 9	260 ± 10
$[0.0, 0.2]$	2780 ± 30	2390 ± 10	2000 ± 20
$[0.2, 0.4]$	13300 ±100	11800 ±100	10200 ±100
$[0.4, 0.6]$	43500 ±300	39500 ±200	35200 ±200
$[0.6, 0.8]$	94000 ±200	88200 ±100	81580 ± 90
$[0.8, 1.0]$	131300 ±300	126700 ±200	121300 ±200
$[1.0, 1.2]$	123800 ±200	122200 ±300	120300 ±300
$[1.2, 1.4]$	90300 ±300	90100 ±300	89800 ±300
$[1.4, 1.6]$	59000 ±300	59000 ±300	59000 ±300
$[1.6, 1.8]$	37700 ±100	37700 ±100	37700 ±100
$[1.8, 2.0]$	24100 ±200	24100 ±200	24100 ±200

Table D.6 Observable binary population: same imagette, $V < 26$, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	289200 ±400	282100 ±400	274100 ±400
$[-0.6, -0.4]$	1.2 ±0.8	1.2 ±0.8	0.8 ±0.5
$[-0.4, -0.2]$	27 ± 2	23 ± 2	19 ± 2
$[-0.2, 0.0]$	242 ± 8	200 ± 10	166 ± 8
$[0.0, 0.2]$	1660 ± 30	1450 ± 30	1230 ± 20
$[0.2, 0.4]$	7450 ± 70	6650 ± 60	5810 ± 60
$[0.4, 0.6]$	22700 ±200	20800 ±100	18800 ±200
$[0.6, 0.8]$	44800 ±200	42500 ±200	39800 ±200
$[0.8, 1.0]$	56300 ±200	54900 ±200	53100 ±200
$[1.0, 1.2]$	49600 ±200	49300 ±200	48900 ±100
$[1.2, 1.4]$	36800 ±100	36700 ±100	36700 ±100
$[1.4, 1.6]$	26300 ±200	26300 ±200	26300 ±200
$[1.6, 1.8]$	19200 ±100	19200 ±100	19200 ±100
$[1.8, 2.0]$	13900 ±100	13900 ±100	13900 ±100

Table D.7 Observable binary population: unblended, all observable by PLATO, $V < 16$, F5-M, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	20600 ±100	20600 ±100	20600 ±100
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	0.5 ±0.6	0.5 ±0.6	0.5 ±0.6
$[0.0,0.2]$	22 ± 6	20 ± 6	18 ± 6
$[0.2,0.4]$	130 ± 7	116 ± 5	102 ± 6
$[0.4,0.6]$	132 ± 6	126 ± 6	115 ± 6
$[0.6,0.8]$	340 ± 20	330 ± 20	320 ± 20
$[0.8,1.0]$	700 ± 20	700 ± 20	690 ± 20
$[1.0,1.2]$	1290 ± 60	1290 ± 60	1290 ± 60
$[1.2,1.4]$	2100 ± 20	2100 ± 20	2100 ± 20
$[1.4,1.6]$	3320 ± 50	3320 ± 50	3320 ± 50
$[1.6,1.8]$	5150 ± 80	5150 ± 80	5150 ± 80
$[1.8,2.0]$	5920 ± 80	5920 ± 80	5920 ± 80

Table D.8 Observable binary population: unblended, $V < 16$, F5-M, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	18300 ±200	18300 ±200	18300±200
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	1 ± 1	0.7 ±0.7	0.7 ±0.7
$[0.0,0.2]$	27 ± 3	24 ± 3	22 ± 3
$[0.2,0.4]$	137 ± 7	127 ± 6	112 ± 4
$[0.4,0.6]$	134 ± 6	129 ± 7	120 ± 5
$[0.6,0.8]$	310 ± 10	300 ± 10	300 ± 10
$[0.8,1.0]$	662 ± 8	659 ± 8	654 ± 7
$[1.0,1.2]$	1160 ± 20	1160 ± 20	1160 ± 20
$[1.2,1.4]$	1920 ± 10	1920 ± 10	1920 ± 10
$[1.4,1.6]$	3100 ± 90	3100 ± 90	3100 ± 90
$[1.6,1.8]$	4700 ±100	4700 ±100	4700 ±100
$[1.8,2.0]$	5050 ± 70	5050 ± 70	5050 ± 70

Table D.9 Observable binary population: same pixel, $V < 16$, F5-M, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	32500 ±200	32100 ±200	31600 ±200
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	1 ± 1	1 ± 1	1 ± 1
$[-0.2,0.0]$	10 ± 4	8 ± 3	7 ± 3
$[0.0,0.2]$	100 ± 20	90 ± 10	80 ± 10
$[0.2,0.4]$	480 ± 10	430 ± 10	370 ± 10
$[0.4,0.6]$	1210 ± 30	1110 ± 30	990 ± 30
$[0.6,0.8]$	2500 ± 20	2370 ± 20	2200 ± 20
$[0.8,1.0]$	3420 ± 60	3350 ± 60	3270 ± 70
$[1.0,1.2]$	3450 ± 70	3440 ± 60	3420 ± 60
$[1.2,1.4]$	3530 ± 20	3530 ± 20	3530 ± 20
$[1.4,1.6]$	4210 ± 70	4210 ± 70	4210 ± 70
$[1.6,1.8]$	5640 ± 80	5640 ± 80	5640 ± 80
$[1.8,2.0]$	6180 ± 70	6180 ± 70	6180 ± 70

Table D.10 Observable binary population: same pixel, $V < 16$, F5-M, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	24100 ±200	23900 ±300	23700 ±300
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0.8 ±0.7	0.7 ±0.8	0.5 ±0.8
$[-0.2,0.0]$	8 ± 2	7 ± 1	7 ± 1
$[0.0,0.2]$	70 ± 8	60 ± 6	54 ± 6
$[0.2,0.4]$	320 ± 20	290 ± 10	257 ± 8
$[0.4,0.6]$	700 ± 9	640 ± 10	582 ± 6
$[0.6,0.8]$	1390 ± 40	1330 ± 40	1260 ± 40
$[0.8,1.0]$	1910 ± 30	1880 ± 30	1850 ± 30
$[1.0,1.2]$	2140 ± 20	2130 ± 20	2130 ± 20
$[1.2,1.4]$	2600 ± 20	2600 ± 20	2600 ± 20
$[1.4,1.6]$	3550 ± 80	3550 ± 80	3550 ± 80
$[1.6,1.8]$	5000 ±100	5000 ±100	5000 ±100
$[1.8,2.0]$	5200 ± 70	5200 ± 70	5200 ± 70

Table D.11 Observable binary population: same imagette, $V < 16$, F5-M, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	443200 ±600	430800 ±500	416800 ±400
$[-0.6, -0.4]$	2 ± 1	1 ± 1	0.9 ±0.5
$[-0.4, -0.2]$	34 ± 2	27 ± 3	21 ± 3
$[-0.2, 0.0]$	370 ± 10	310 ± 10	250 ± 10
$[0.0, 0.2]$	2610 ± 30	2240 ± 20	1880 ± 20
$[0.2, 0.4]$	12100 ±100	10800 ±100	9300 ±100
$[0.4, 0.6]$	38200 ±300	34800 ±200	30200 ±200
$[0.6, 0.8]$	78300 ±200	73900 ±200	68800 ±200
$[0.8, 1.0]$	96700 ±200	94400 ±200	91600 ±200
$[1.0, 1.2]$	78700 ±200	78200 ±200	77700 ±200
$[1.2, 1.4]$	53800 ±200	53700 ±200	53700 ±200
$[1.4, 1.6]$	34800 ±100	34800 ±100	34800 ±100
$[1.6, 1.8]$	22900 ±200	22900 ±200	22900 ±200
$[1.8, 2.0]$	15200 ±200	15200 ±200	15200 ±200

Table D.12 Observable binary population: same imagette, $V < 16$, F5-M, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	226000 ±400	220600 ±400	214400 ±300
$[-0.6, -0.4]$	1.1 ±0.7	1.1 ±0.7	0.7 ±0.4
$[-0.4, -0.2]$	26 ± 2	23 ± 2	18 ± 2
$[-0.2, 0.0]$	230 ± 8	190 ± 10	157 ± 7
$[0.0, 0.2]$	1560 ± 30	1370 ± 40	1160 ± 30
$[0.2, 0.4]$	6810 ± 70	6090 ± 50	5330 ± 60
$[0.4, 0.6]$	20200 ±200	18600 ±200	16800 ±200
$[0.6, 0.8]$	38500 ±200	36600 ±100	34400 ±100
$[0.8, 1.0]$	44800 ±200	44000 ±200	42900 ±200
$[1.0, 1.2]$	36800 ±200	36700 ±200	36500 ±200
$[1.2, 1.4]$	26610 ± 80	26600 ± 80	26590 ± 80
$[1.4, 1.6]$	19200 ±200	19200 ±200	19200 ±200
$[1.6, 1.8]$	14400 ±100	14400 ±100	14400 ±100
$[1.8, 2.0]$	10500 ±100	10500 ±100	10500 ±100

Table D.13 Observable binary population: unblended, P1, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	134 ±9	134 ±9	134 ±9
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	0.2 ±0.5	0.2 ±0.5	0.2 ±0.5
$[0.0,0.2]$	1.0 ±0.7	1.0 ±0.7	1.0 ±0.7
$[0.2,0.4]$	3 ±1	3 ±1	3 ±1
$[0.4,0.6]$	1 ±1	1 ±1	1 ±1
$[0.6,0.8]$	2.1 ±0.7	2.1 ±0.7	2.1 ±0.7
$[0.8,1.0]$	3 ±1	3 ±1	3 ±1
$[1.0,1.2]$	7 ±2	7 ±2	7 ±2
$[1.2,1.4]$	10 ±2	10 ±2	10 ±2
$[1.4,1.6]$	17 ±4	17 ±4	17 ±4
$[1.6,1.8]$	27 ±6	27 ±6	27 ±6
$[1.8,2.0]$	41 ±2	41 ±2	41 ±2

Table D.14 Observable binary population: unblended, P1, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	158 ±7	158 ±7	158 ±7
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	0.1 ±0.2	0	0
$[0.0,0.2]$	3 ±1	3 ±1	3 ±1
$[0.2,0.4]$	4 ±1	4 ±1	4 ±1
$[0.4,0.6]$	1 ±1	1 ±1	1 ±1
$[0.6,0.8]$	1 ±1	1 ±1	1 ±1
$[0.8,1.0]$	4 ±2	4 ±2	4 ±2
$[1.0,1.2]$	9 ±3	9 ±3	9 ±3
$[1.2,1.4]$	11 ±3	11 ±3	11 ±3
$[1.4,1.6]$	19 ±5	19 ±5	19 ±5
$[1.6,1.8]$	30 ±5	30 ±5	30 ±5
$[1.8,2.0]$	46 ±4	46 ±4	46 ±4

Table D.15 Observable binary population: same pixel, P1, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	158 ±7	157 ±8	157 ±8
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0.5 ±0.6	0.5 ±0.6	0.3 ±0.5
$[-0.2,0.0]$	3 ±2	2 ±1	2 ±1
$[0.0,0.2]$	8 ±1	7 ±1	7 ±1
$[0.2,0.4]$	9 ±1	9 ±1	9 ±1
$[0.4,0.6]$	5 ±2	5 ±2	5 ±2
$[0.6,0.8]$	4 ±2	4 ±2	4 ±2
$[0.8,1.0]$	5 ±2	5 ±2	5 ±2
$[1.0,1.2]$	9 ±2	9 ±2	9 ±2
$[1.2,1.4]$	11 ±2	11 ±2	11 ±2
$[1.4,1.6]$	17 ±3	17 ±3	17 ±3
$[1.6,1.8]$	27 ±6	27 ±6	27 ±6
$[1.8,2.0]$	41 ±2	41 ±2	41 ±2

Table D.16 Observable binary population: same pixel, P1, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	174 ±5	174 ±5	173 ±6
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0.4 ±0.5	0.4 ±0.5	0.2 ±0.4
$[-0.2,0.0]$	1 ±1	1 ±1	1.1 ±0.7
$[0.0,0.2]$	5 ±2	5 ±2	5 ±2
$[0.2,0.4]$	8 ±1	8 ±1	7.9 ±0.9
$[0.4,0.6]$	4 ±1	4 ±1	4 ±1
$[0.6,0.8]$	4 ±2	4 ±2	4 ±2
$[0.8,1.0]$	6 ±1	6 ±1	6 ±1
$[1.0,1.2]$	9 ±3	9 ±3	9 ±3
$[1.2,1.4]$	11 ±3	11 ±3	11 ±3
$[1.4,1.6]$	19 ±5	19 ±5	19 ±5
$[1.6,1.8]$	30 ±5	30 ±5	30 ±5
$[1.8,2.0]$	46 ±4	46 ±4	46 ±4

Table D.17 Observable binary population: same imagette, P1, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	1010 ±20	980 ±20	950 ±10
$[-0.6,-0.4]$	0.8 ±0.9	0.7 ±0.9	0.4 ±0.5
$[-0.4,-0.2]$	10 ± 2	9 ± 1	7 ± 1
$[-0.2,0.0]$	71 ± 8	63 ± 7	54 ± 6
$[0.0,0.2]$	190 ±10	180 ±10	169 ± 9
$[0.2,0.4]$	230 ±10	220 ±10	220 ±10
$[0.4,0.6]$	164 ± 9	160 ±10	162 ± 9
$[0.6,0.8]$	99 ± 7	99 ± 7	99 ± 7
$[0.8,1.0]$	58 ± 3	58 ± 3	58 ± 3
$[1.0,1.2]$	30 ± 3	30 ± 3	30 ± 3
$[1.2,1.4]$	28 ± 4	28 ± 4	28 ± 4
$[1.4,1.6]$	26 ± 4	26 ± 4	26 ± 4
$[1.6,1.8]$	32 ± 7	32 ± 7	32 ± 7
$[1.8,2.0]$	42 ± 2	42 ± 2	42 ± 2

Table D.18 Observable binary population: same imagette, P1, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	750 ±10	730 ±10	720 ±10
$[-0.6,-0.4]$	0.1 ±0.1	0.1 ±0.1	0.1 ±0.1
$[-0.4,-0.2]$	9 ± 2	9 ± 2	8 ± 2
$[-0.2,0.0]$	51 ± 6	44 ± 6	38 ± 6
$[0.0,0.2]$	130 ±10	120 ±10	111 ± 9
$[0.2,0.4]$	140 ±10	140 ±10	130 ±10
$[0.4,0.6]$	110 ±10	110 ±10	110 ±10
$[0.6,0.8]$	76 ± 6	76 ± 6	76 ± 6
$[0.8,1.0]$	46 ± 4	46 ± 4	46 ± 4
$[1.0,1.2]$	35 ± 9	35 ± 9	35 ± 9
$[1.2,1.4]$	23 ± 8	23 ± 8	23 ± 8
$[1.4,1.6]$	25 ± 3	25 ± 3	25 ± 3
$[1.6,1.8]$	33 ± 6	33 ± 6	33 ± 6
$[1.8,2.0]$	48 ± 5	48 ± 5	48 ± 5

Table D.19 Observable binary population: unblended, P2, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	1.4 ±0.8	1.4 ±0.8	1.4 ±0.8
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	0	0	0
$[0.0,0.2]$	0	0	0
$[0.2,0.4]$	0	0	0
$[0.4,0.6]$	0	0	0
$[0.6,0.8]$	0	0	0
$[0.8,1.0]$	0	0	0
$[1.0,1.2]$	0.01 ±0.03	0.01 ±0.03	0.01 ±0.03
$[1.2,1.4]$	0.1 ±0.1	0.1 ±0.1	0.1 ±0.1
$[1.4,1.6]$	0.02 ±0.02	0.02 ±0.02	0.02 ±0.02
$[1.6,1.8]$	0.7 ±0.8	0.7 ±0.8	0.7 ±0.8
$[1.8,2.0]$	0.3 ±0.5	0.3 ±0.5	0.3 ±0.5

Table D.20 Observable binary population: unblended, P2, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	2.3 ±0.9	2.3 ±0.9	2.3 ±0.9
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	0	0	0
$[0.0,0.2]$	0.2 ±0.2	0.2 ±0.2	0.2 ±0.2
$[0.2,0.4]$	0.2 ±0.4	0.2 ±0.4	0.2 ±0.4
$[0.4,0.6]$	0.001 ±0.002	0.001 ±0.002	0.001 ±0.002
$[0.6,0.8]$	0	0	0
$[0.8,1.0]$	0.1 ±0.2	0.1 ±0.2	0.1 ±0.2
$[1.0,1.2]$	0.4 ±0.4	0.4 ±0.4	0.4 ±0.4
$[1.2,1.4]$	0.04 ±0.07	0.04 ±0.07	0.04 ±0.07
$[1.4,1.6]$	0.2 ±0.4	0.2 ±0.4	0.2 ±0.4
$[1.6,1.8]$	0.2 ±0.2	0.2 ±0.2	0.2 ±0.2
$[1.8,2.0]$	0.5 ±0.5	0.5 ±0.5	0.5 ±0.5

Table D.21 Observable binary population: same pixel, P2, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	1.6 ±0.5	1.6 ±0.5	1.4 ±0.8
$[-0.6, -0.4]$	0	0	0
$[-0.4, -0.2]$	0	0	0
$[-0.2, 0.0]$	0	0	0
$[0.0, 0.2]$	0	0	0
$[0.2, 0.4]$	0	0	0
$[0.4, 0.6]$	0	0	0
$[0.6, 0.8]$	0	0	0
$[0.8, 1.0]$	0	0	0
$[1.0, 1.2]$	0.01 ±0.03	0.01 ±0.03	0.01 ±0.03
$[1.2, 1.4]$	0.1 ±0.1	0.1 ±0.1	0.1 ±0.1
$[1.4, 1.6]$	0.02 ±0.02	0.02 ±0.02	0.02 ±0.02
$[1.6, 1.8]$	0.7 ±0.8	0.7 ±0.8	0.7 ±0.8
$[1.8, 2.0]$	0.3 ±0.5	0.3 ±0.5	0.3 ±0.5

Table D.22 Observable binary population: same pixel, P2, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	2.3 ±0.9	2.3 ±0.9	2.3 ±0.9
$[-0.6, -0.4]$	0	0	0
$[-0.4, -0.2]$	0	0	0
$[-0.2, 0.0]$	0	0	0
$[0.0, 0.2]$	0.2 ±0.2	0.2 ±0.2	0.2 ±0.2
$[0.2, 0.4]$	0.2 ±0.4	0.2 ±0.4	0.2 ±0.4
$[0.4, 0.6]$	0.001 ±0.002	0.001 ±0.002	0.001 ±0.002
$[0.6, 0.8]$	0	0	0
$[0.8, 1.0]$	0.1 ±0.2	0.1 ±0.2	0.1 ±0.2
$[1.0, 1.2]$	0.4 ±0.4	0.4 ±0.4	0.4 ±0.4
$[1.2, 1.4]$	0.04 ±0.07	0.04 ±0.07	0.04 ±0.07
$[1.4, 1.6]$	0.2 ±0.4	0.2 ±0.4	0.2 ±0.4
$[1.6, 1.8]$	0.2 ±0.2	0.2 ±0.2	0.2 ±0.2
$[1.8, 2.0]$	0.5 ±0.5	0.5 ±0.5	0.5 ±0.5

Table D.23 Observable binary population: same imagette, P2, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	7 ±2	6 ±2	6 ±2
$[-0.6,-0.4]$	0.2 ±0.4	0.2 ±0.4	0.2 ±0.4
$[-0.4,-0.2]$	1 ±1	1 ±1	1 ±1
$[-0.2,0.0]$	1.0 ±0.7	0.8 ±0.4	0.8 ±0.4
$[0.0,0.2]$	0.4 ±0.8	0.2 ±0.4	0.2 ±0.4
$[0.2,0.4]$	1 ±2	1 ±2	1 ±2
$[0.4,0.6]$	0.8 ±0.7	0.8 ±0.7	0.8 ±0.7
$[0.6,0.8]$	0	0	0
$[0.8,1.0]$	0.2 ±0.4	0.2 ±0.4	0.2 ±0.4
$[1.0,1.2]$	0.01 ±0.03	0.01 ±0.03	0.01 ±0.03
$[1.2,1.4]$	0.1 ±0.1	0.1 ±0.1	0.1 ±0.1
$[1.4,1.6]$	0.02 ±0.02	0.02 ±0.02	0.02 ±0.02
$[1.6,1.8]$	0.7 ±0.8	0.7 ±0.8	0.7 ±0.8
$[1.8,2.0]$	0.3 ±0.5	0.3 ±0.5	0.3 ±0.5

Table D.24 Observable binary population: same imagette, P2, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	6.7 ±0.8	6.7 ±0.8	6.7 ±0.8
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	1.0 ±0.7	0.9 ±0.7	0.9 ±0.7
$[-0.2,0.0]$	1.0 ±0.9	1.0 ±0.9	1.0 ±0.9
$[0.0,0.2]$	0.4 ±0.3	0.4 ±0.3	0.4 ±0.4
$[0.2,0.4]$	1 ±1	1 ±1	1 ±1
$[0.4,0.6]$	1.3 ±0.4	1.3 ±0.4	1.3 ±0.4
$[0.6,0.8]$	0.2 ±0.4	0.2 ±0.4	0.2 ±0.4
$[0.8,1.0]$	0.4 ±0.4	0.4 ±0.4	0.4 ±0.4
$[1.0,1.2]$	0.4 ±0.4	0.4 ±0.4	0.4 ±0.4
$[1.2,1.4]$	0.04 ±0.07	0.04 ±0.07	0.04 ±0.07
$[1.4,1.6]$	0.2 ±0.4	0.2 ±0.4	0.2 ±0.4
$[1.6,1.8]$	0.2 ±0.2	0.2 ±0.2	0.2 ±0.2
$[1.8,2.0]$	0.5 ±0.5	0.5 ±0.5	0.5 ±0.5

Table D.25 Observable binary population: unblended, P4, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	190 ±10	190 ±10	180 ±10
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	0	0	0
$[0.0,0.2]$	11 ± 3	10 ± 3	9 ± 3
$[0.2,0.4]$	29 ± 3	26 ± 3	23 ± 3
$[0.4,0.6]$	9 ± 2	9 ± 2	9 ± 2
$[0.6,0.8]$	9 ± 1	9 ± 1	9 ± 1
$[0.8,1.0]$	15 ± 5	15 ± 5	15 ± 5
$[1.0,1.2]$	26 ± 3	26 ± 3	26 ± 3
$[1.2,1.4]$	49 ± 7	49 ± 7	49 ± 7
$[1.4,1.6]$	44 ± 5	44 ± 5	44 ± 5
$[1.6,1.8]$	0.2 ±0.3	0.2 ±0.3	0.2 ±0.3
$[1.8,2.0]$	0	0	0

Table D.26 Observable binary population: unblended, P4, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	220 ± 8	218 ± 8	214 ± 9
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	0.4 ±0.6	0.4 ±0.6	0.4 ±0.5
$[0.0,0.2]$	11 ± 3	11 ± 3	9 ± 3
$[0.2,0.4]$	30 ± 2	29 ± 1	27 ± 2
$[0.4,0.6]$	8 ± 2	8 ± 2	7 ± 1
$[0.6,0.8]$	10 ± 2	10 ± 2	10 ± 2
$[0.8,1.0]$	18 ± 2	18 ± 2	18 ± 2
$[1.0,1.2]$	29 ± 6	29 ± 6	29 ± 6
$[1.2,1.4]$	60 ±10	60 ±10	60 ±10
$[1.4,1.6]$	55 ± 6	55 ± 6	55 ± 6
$[1.6,1.8]$	0.5 ±0.6	0.5 ±0.6	0.5 ±0.6
$[1.8,2.0]$	0	0	0

Table D.27 Observable binary population: same pixel, P4, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	550 ±20	540 ±20	520 ±20
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0.4 ±0.6	0.4 ±0.6	0.4 ±0.6
$[-0.2,0.0]$	2 ± 1	1.8 ±0.9	1.6 ±0.8
$[0.0,0.2]$	28 ± 7	25 ± 6	22 ± 5
$[0.2,0.4]$	67 ± 7	63 ± 7	58 ± 7
$[0.4,0.6]$	69 ± 6	67 ± 6	63 ± 7
$[0.6,0.8]$	80 ± 7	78 ± 7	75 ± 5
$[0.8,1.0]$	80 ±20	80 ±20	80 ±20
$[1.0,1.2]$	69 ± 4	69 ± 4	69 ± 4
$[1.2,1.4]$	75 ± 6	75 ± 6	75 ± 6
$[1.4,1.6]$	60 ± 2	60 ± 2	60 ± 2
$[1.6,1.8]$	10 ± 2	10 ± 2	10 ± 2
$[1.8,2.0]$	4 ± 3	4 ± 3	4 ± 3

Table D.28 Observable binary population: same pixel, P4, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	413 ± 9	410 ±10	400 ± 9
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0.3 ±0.4	0.2 ±0.4	0.2 ±0.4
$[-0.2,0.0]$	3 ± 1	2.7 ±0.9	2.7 ±0.9
$[0.0,0.2]$	21 ± 3	19 ± 4	17 ± 3
$[0.2,0.4]$	51 ± 5	48 ± 5	45 ± 5
$[0.4,0.6]$	42 ± 6	41 ± 5	38 ± 7
$[0.6,0.8]$	47 ± 3	45 ± 3	44 ± 3
$[0.8,1.0]$	51 ± 4	51 ± 4	51 ± 4
$[1.0,1.2]$	52 ± 5	52 ± 5	52 ± 5
$[1.2,1.4]$	70 ±10	70 ±10	70 ±10
$[1.4,1.6]$	65 ± 5	65 ± 5	65 ± 5
$[1.6,1.8]$	6 ± 1	6 ± 1	6 ± 1
$[1.8,2.0]$	3 ± 1	3 ± 1	3 ± 1

Table D.29 Observable binary population: same imagette, P4, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	12800 ±200	12500 ±200	12100 ±100
$[-0.6,-0.4]$	0.9 ±0.5	0.6 ±0.4	0.5 ±0.4
$[-0.4,-0.2]$	18 ± 3	15 ± 3	12 ± 3
$[-0.2,0.0]$	134 ± 8	117 ± 9	100 ± 9
$[0.0,0.2]$	560 ± 30	500 ± 20	440 ± 30
$[0.2,0.4]$	1290 ± 30	1210 ± 20	1120 ± 20
$[0.4,0.6]$	2100 ± 40	1990 ± 50	1870 ± 50
$[0.6,0.8]$	2650 ± 70	2580 ± 70	2490 ± 50
$[0.8,1.0]$	2280 ± 40	2270 ± 40	2260 ± 40
$[1.0,1.2]$	1510 ± 50	1510 ± 50	1510 ± 50
$[1.2,1.4]$	1031 ± 40	1031 ± 40	1031 ± 40
$[1.4,1.6]$	620 ± 20	620 ± 20	620 ± 20
$[1.6,1.8]$	330 ± 20	330 ± 20	330 ± 20
$[1.8,2.0]$	160 ± 10	160 ± 10	160 ± 10

Table D.30 Observable binary population: same imagette, P4, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	6940 ±50	6780 ±50	6580 ±50
$[-0.6,-0.4]$	1.0 ±0.7	1.0 ±0.6	0.7 ±0.3
$[-0.4,-0.2]$	13 ± 3	12 ± 2	9 ± 2
$[-0.2,0.0]$	78 ± 4	67 ± 4	57 ± 4
$[0.0,0.2]$	320 ±20	290 ±10	270 ±20
$[0.2,0.4]$	740 ±20	690 ±20	640 ±20
$[0.4,0.6]$	1090 ±30	1050 ±30	990 ±30
$[0.6,0.8]$	1350 ±10	1320 ±10	1280 ±10
$[0.8,1.0]$	1120 ±30	1120 ±30	1110 ±30
$[1.0,1.2]$	840 ±20	840 ±20	840 ±20
$[1.2,1.4]$	560 ±20	560 ±20	560 ±20
$[1.4,1.6]$	400 ±10	400 ±10	400 ±10
$[1.6,1.8]$	210 ±10	210 ±10	210 ±10
$[1.8,2.0]$	110 ±10	110 ±10	110 ±10

Table D.31 Observable binary population: unblended, P5, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	1350 ±30	1340 ±30	1340 ±30
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	0.5 ±0.6	0.5 ±0.6	0.5 ±0.6
$[0.0,0.2]$	9 ± 4	8 ± 4	7 ± 4
$[0.2,0.4]$	35 ± 2	33 ± 2	32 ± 3
$[0.4,0.6]$	16 ± 3	16 ± 3	15 ± 3
$[0.6,0.8]$	21 ± 4	21 ± 4	21 ± 4
$[0.8,1.0]$	36 ± 1	36 ± 1	36 ± 1
$[1.0,1.2]$	67 ± 9	67 ± 9	67 ± 9
$[1.2,1.4]$	120 ±10	120 ±10	120 ±10
$[1.4,1.6]$	180 ±10	180 ±10	180 ±10
$[1.6,1.8]$	290 ±10	290 ±10	290 ±10
$[1.8,2.0]$	400 ±20	400 ±20	400 ±20

Table D.32 Observable binary population: unblended, P5, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	1500 ±50	1500 ±50	1500 ±40
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0	0	0
$[-0.2,0.0]$	0.5 ±0.8	0.3 ±0.4	0.3 ±0.4
$[0.0,0.2]$	11 ± 3	10 ± 2	9 ± 3
$[0.2,0.4]$	42 ± 6	41 ± 6	38 ± 4
$[0.4,0.6]$	17 ± 3	17 ± 3	17 ± 3
$[0.6,0.8]$	28 ± 4	28 ± 4	28 ± 4
$[0.8,1.0]$	48 ± 4	48 ± 4	48 ± 4
$[1.0,1.2]$	73 ± 4	73 ± 4	73 ± 4
$[1.2,1.4]$	121 ± 9	121 ± 9	121 ± 9
$[1.4,1.6]$	210 ±20	210 ±20	210 ±20
$[1.6,1.8]$	330 ±20	330 ±20	330 ±20
$[1.8,2.0]$	450 ±10	450 ±10	450 ±10

Table D.33 Observable binary population: same pixel, P5, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	1738 ±30	1725 ±30	1707 ±30
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0.6 ±0.5	0.6 ±0.5	0.3 ±0.5
$[-0.2,0.0]$	7 ± 3	6 ± 2	5 ± 2
$[0.0,0.2]$	43 ± 8	39 ± 9	34 ± 7
$[0.2,0.4]$	121 ± 7	115 ± 7	107 ± 8
$[0.4,0.6]$	110 ±10	110 ±10	110 ±10
$[0.6,0.8]$	90 ± 5	90 ± 5	89 ± 4
$[0.8,1.0]$	84 ± 5	84 ± 5	84 ± 5
$[1.0,1.2]$	90 ±10	90 ±10	90 ±10
$[1.2,1.4]$	134 ± 9	134 ± 9	134 ± 9
$[1.4,1.6]$	190 ±10	190 ±10	190 ±10
$[1.6,1.8]$	290 ±10	290 ±10	290 ±10
$[1.8,2.0]$	400 ±20	400 ±20	400 ±20

Table D.34 Observable binary population: same pixel, P5, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	1760 ±50	1750 ±50	1740 ±50
$[-0.6,-0.4]$	0	0	0
$[-0.4,-0.2]$	0.5 ±0.4	0.5 ±0.5	0.2 ±0.4
$[-0.2,0.0]$	5 ± 2	4 ± 2	4 ± 1
$[0.0,0.2]$	32 ± 4	28 ± 3	26 ± 3
$[0.2,0.4]$	91 ± 9	88 ± 7	82 ± 6
$[0.4,0.6]$	75 ± 8	74 ± 8	73 ± 9
$[0.6,0.8]$	76 ± 7	76 ± 7	75 ± 7
$[0.8,1.0]$	80 ± 5	80 ± 5	80 ± 5
$[1.0,1.2]$	91 ± 2	91 ± 2	91 ± 2
$[1.2,1.4]$	130 ±10	130 ±10	130 ±10
$[1.4,1.6]$	220 ±20	220 ±20	220 ±20
$[1.6,1.8]$	330 ±20	330 ±20	330 ±20
$[1.8,2.0]$	450 ±10	450 ±10	450 ±10

Table D.35 Observable binary population: same imagette, P5, LLN

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	15240 ±50	14820 ±70	14340 ±60
$[-0.6,-0.4]$	0.9 ±0.9	0.7 ±0.9	0.4 ±0.5
$[-0.4,-0.2]$	16 ± 1	13 ± 2	10 ± 1
$[-0.2,0.0]$	200 ±10	167 ± 9	140 ±50
$[0.0,0.2]$	1180 ±10	1050 ±20	900 ±20
$[0.2,0.4]$	2940 ±30	2770 ±30	2570 ±30
$[0.4,0.6]$	3370 ±30	3310 ±30	3220 ±30
$[0.6,0.8]$	2570 ±30	2560 ±30	2540 ±30
$[0.8,1.0]$	1700 ±40	1700 ±40	1700 ±40
$[1.0,1.2]$	1040 ±20	1040 ±20	1040 ±20
$[1.2,1.4]$	660 ±30	660 ±30	660 ±30
$[1.4,1.6]$	460 ±30	460 ±30	460 ±30
$[1.6,1.8]$	420 ±10	420 ±10	420 ±10
$[1.8,2.0]$	470 ±30	470 ±30	470 ±30

Table D.36 Observable binary population: same imagette, P5, LLS

Apparent radius/log $R_\oplus$	32 cameras	28 cameras	24 cameras
All	6940 ±50	6780 ±50	6580 ±50
$[-0.6,-0.4]$	0.1 ±0.1	0.1 ±0.1	0.1 ±0.1
$[-0.4,-0.2]$	12 ± 4	11 ± 4	9 ± 2
$[-0.2,0.0]$	130 ±10	110 ±10	90 ±10
$[0.0,0.2]$	730 ±30	660 ±30	580 ±20
$[0.2,0.4]$	1730 ±30	1630 ±30	1520 ±20
$[0.4,0.6]$	2000 ±60	1960 ±60	1920 ±60
$[0.6,0.8]$	1610 ±30	1610 ±30	1600 ±30
$[0.8,1.0]$	1140 ±20	1140 ±20	1140 ±20
$[1.0,1.2]$	740 ±20	740 ±20	740 ±20
$[1.2,1.4]$	530 ±20	530 ±20	530 ±20
$[1.4,1.6]$	420 ±20	420 ±20	420 ±20
$[1.6,1.8]$	440 ±10	440 ±10	440 ±10
$[1.8,2.0]$	510 ±10	510 ±10	510 ±10

Appendix E

Detailed results of planet to binary ratios

Table E.1 Detailed results, planets to binaries, LLN, unblended binaries, true radius of planets, 24 cameras. $\delta X/X$ is the fractional uncertainty. NB indicates no binaries, NP indicates no planets. For use with Fig. 8.2, upper panel.

Radius/ *log* $R_\oplus$	$V<26$	$\delta X/X$	$V<16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	NB		NB		NB		NB		NB		NB	
-0.2 to 0	2400	1.20	2200	1.20	1700	2.50	NB		NB		1700	1.20
0 to 0.2	240	0.25	230	0.33	730	0.70	NB		48	0.33	360	0.57
0.2 to 0.4	64	0.06	61	0.06	130	0.34	NB		12	0.15	78	0.10
0.4 to 0.6	37	0.05	39	0.06	77	1.00	NB		7.9	0.26	57	0.21
0.6 to 0.8	4.5	0.05	3.6	0.06	5.6	0.47	NB		0.87	0.39	5.4	0.21
0.8 to 1.0	1.7	0.03	0.97	0.04	1.8	0.60	NB		0.19	0.75	1.3	0.15
1.0 to 1.2	1.4	0.05	0.50	0.06	0.62	0.66	22	3.43	0.15	0.51	0.60	0.18
1.2 to 1.4	0.16	0.05	0.033	0.09	0.029	5.27	NP		0.0020	3.00	0.033	0.51
'omega'												
-0.4 to -0.2	NB		NB		NB		NB		NB		NB	
-0.2 to 0	170	1.20	160	1.20	120	2.50	NB		NB		120	1.20
0 to 0.2	26	0.25	26	0.34	74	0.70	NB		5.1	1.62	74	0.57
0.2 to 0.4	14	0.06	14	0.06	28	0.34	NB		2.7	0.16	28	0.10
0.4 to 0.6	18	0.05	19	0.05	36	1.00	NB		3.7	0.26	36	0.20
0.6 to 0.8	3.4	0.05	2.7	0.06	3.7	0.38	NB		0.62	0.22	3.7	0.20
0.8 to 1.0	1.3	0.03	0.75	0.03	1.1	0.40	NB		0.12	0.69	1.1	0.07
1.0 to 1.2	1.2	0.04	0.43	0.07	0.47	0.61	7.3	3.61	0.11	0.49	0.47	0.15
1.2 to 1.4	0.19	0.02	0.041	0.07	0.020	5.14	NP		0.014	0.74	0.020	0.21

Table E.2 As Table E.1, LLN, unblended binaries, apparent radius of planets after blending within the pixel. For use with Fig. 8.2, lower panel.

Radius/ *log* $R_\oplus$	$V<26$	$\delta X/X$	$V<16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	NB		NB		NB		NB		NB		NB	
-0.2 to 0	2400	1.20	2200	1.20	1900	2.50	NB		NB		1800	1.20
0 to 0.2	220	0.25	220	0.33	710	0.70	NB		40	0.33	350	0.57
0.2 to 0.4	59	0.06	57	0.06	130	0.34	NB		10	0.14	76	0.10
0.4 to 0.6	30	0.05	33	0.06	73	1.00	NB		6.4	0.28	53	0.21
0.6 to 0.8	3.6	0.07	3.1	0.07	5.6	0.47	NB		0.81	0.27	5.3	0.21
0.8 to 1.0	1.4	0.03	0.93	0.04	1.9	0.60	NB		0.22	0.75	1.3	0.15
1.0 to 1.2	0.53	0.06	0.36	0.06	0.60	0.58	22	3.43	0.11	0.68	0.56	0.19
1.2 to 1.4	0.036	0.11	0.025	0.11	0.029	5.27	NP		0.0020	3.00	0.030	0.51
'omega'												
-0.4 to -0.2	NB		NB		NB		NB		NB		NB	
-0.2 to 0	180	1.20	170	1.20	130	2.50	NB		NB		130	1.20
0 to 0.2	25	0.25	25	0.34	72	0.70	NB		4.6	0.39	39	0.57
0.2 to 0.4	14	0.06	13	0.07	29	0.35	NB		2.4	0.16	17	0.10
0.4 to 0.6	15	0.05	16	0.05	35	1.00	NB		3.0	0.28	26	0.20
0.6 to 0.8	2.7	0.05	2.4	0.06	3.7	0.37	NB		0.56	0.22	4.0	0.20
0.8 to 1.0	1.1	0.03	0.72	0.04	1.1	0.39	NB		0.13	0.46	1.0	0.08
1.0 to 1.2	0.47	0.05	0.32	0.06	0.46	0.66	7.3	3.61	0.11	0.35	0.46	0.16
1.2 to 1.4	0.044	0.05	0.031	0.06	0.020	5.14	NP		0.0082	1.01	0.036	0.15

Table E.3 As Table E.1, LLS, unblended binaries, true radius of planets. For use with Fig. 8.3, upper panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	NB		NB		NB		NB		NB		NB	
-0.2 to 0	1900	1.00	1800	1.00	NB		NB		550	1.25	3100	1.33
0 to 0.2	220	0.13	220	0.14	280	0.34	130	1.02	49	0.34	330	0.33
0.2 to 0.4	69	0.05	62	0.04	100	0.26	39	2.01	11	0.10	75	0.11
0.4 to 0.6	35	0.03	39	0.04	100	1.00	2000	2.24	10	0.17	57	0.18
0.6 to 0.8	4.8	0.05	3.9	0.05	9.0	1.05	NB		0.97	0.36	4.6	0.16
0.8 to 1.0	1.6	0.02	0.94	0.02	1.2	0.64	2.6	2.83	0.21	0.51	1.0	0.15
1.0 to 1.2	1.1	0.02	0.51	0.04	0.53	0.39	NP		0.13	0.32	0.63	0.14
1.2 to 1.4	0.10	0.05	0.031	0.13	0.056	1.20	NP		0.0035	2.01	0.026	0.51
'omega'												
-0.4 to -0.2	NB		NB		NB		NB		NB		NB	
-0.2 to 0	130	1.00	130	1.00	NB		NB		35	1.25	223	1.34
0 to 0.2	25	0.13	24	0.14	31	0.34	11	1.14	5.1	0.35	36	0.33
0.2 to 0.4	15	0.05	14	0.04	23	0.27	4.6	2.12	2.5	0.08	17	0.11
0.4 to 0.6	17	0.03	19	0.04	48	1.00	600	2.60	5.2	0.18	27	0.18
0.6 to 0.8	3.7	0.03	2.9	0.04	6.8	1.02	NB		0.67	0.39	3.8	0.14
0.8 to 1.0	1.2	0.02	0.72	0.03	1.1	0.52	NP		0.16	0.31	0.79	0.10
1.0 to 1.2	0.95	0.02	0.43	0.03	0.53	0.52	0.28	2.24	0.075	0.34	0.56	0.09
1.2 to 1.4	0.12	0.04	0.038	0.03	0.065	0.63	NP		0.0052	1.34	0.056	0.23

Table E.4 As Table E.1, LLS, unblended binaries, apparent radius of planets after blending within the pixel. For use with Fig. 8.3, lower panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	NB		NB		NB		NB		NB		NB	
-0.2 to 0	2000	1.00	1800	1.00	NB		NB		530	1.25	3200	1.33
0 to 0.2	210	0.13	210	0.14	270	0.34	120	1.02	43	0.33	310	0.33
0.2 to 0.4	65	0.05	59	0.04	110	0.26	39	2.01	9.9	0.10	74	0.11
0.4 to 0.6	30	0.03	34	0.04	96	1.01	2000	2.24	8.8	0.16	53	0.18
0.6 to 0.8	4.1	0.05	3.5	0.05	9.0	1.05	NB		0.91	0.39	4.6	0.16
0.8 to 1.0	1.3	0.03	0.92	0.04	1.2	0.64	2.6	2.83	0.18	0.68	1.1	0.13
1.0 to 1.2	0.57	0.02	0.41	0.05	0.52	0.39	NP		0.13	0.32	0.52	0.15
1.2 to 1.4	0.037	0.12	0.026	0.12	0.056	1.20	NP		0.0035	2.01	0.025	0.51
'omega'												
-0.4 to -0.2	NB		NB		NB		NB		NB		NB	
-0.2 to 0	140	1.00	130	1.00	NB		NB		34	1.25	240	1.34
0 to 0.2	24	0.13	23	0.14	31	0.34	11	1.18	4.9	0.34	35	0.33
0.2 to 0.4	15	0.05	14	0.04	23	0.26	4.6	2.12	2.4	0.08	17	0.11
0.4 to 0.6	15	0.03	17	0.04	46	1.00	600	2.60	4.5	0.18	26	0.18
0.6 to 0.8	3.1	0.04	2.7	0.04	6.7	1.02	NB		0.71	0.32	3.7	0.15
0.8 to 1.0	1.0	0.03	0.72	0.04	1.1	0.52	NP		0.14	0.47	0.84	0.09
1.0 to 1.2	0.49	0.02	0.35	0.03	0.50	0.60	2.8	2.24	0.068	0.34	0.52	0.12
1.2 to 1.4	0.043	0.05	0.031	0.03	0.065	0.63	NP		0.0052	1.34	0.054	0.23

Table E.5 As Table E.1, LLN, binaries within 7.5", true radius of planets. For use with Fig. 8.4, upper panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	140	1.00	140	1.00	200	1.67	NB		69	1.51	190	1.67
-0.2 to 0	170	0.29	160	0.43	200	0.50	NB		110	0.50	170	0.40
0 to 0.2	57	0.13	55	0.13	100	0.15	NB		18	0.23	78	0.21
0.2 to 0.4	18	0.03	17	0.03	43	0.12	NB		4.7	0.14	23	0.08
0.4 to 0.6	5.1	0.03	4.6	0.04	21	0.40	NB		1.1	0.17	8.3	0.10
0.6 to 0.8	0.66	0.02	0.53	0.02	3.1	0.60	NB		0.11	0.38	1.3	0.10
0.8 to 1.0	0.32	0.03	0.21	0.04	1.1	0.64	NB		0.035	0.71	0.56	0.16
1.0 to 1.2	0.42	0.03	0.19	0.04	0.53	0.64	22	3.43	0.055	0.50	0.44	0.17
1.2 to 1.4	0.078	0.05	0.020	0.09	0.028	1.68	NP		0.0013	3.00	0.028	0.50
'omega'												
-0.4 to -0.2	6.6	1.03	6.2	1.06	12	1.71	NB		2.2	1.77	9.5	1.72
-0.2 to 0	12	0.29	12	0.43	14	0.51	NB		9.3	0.55	12	0.40
0 to 0.2	6.3	0.13	6.1	0.13	11	0.15	NB		2.0	1.60	8.6	0.21
0.2 to 0.4	4.0	0.03	3.7	0.03	9.5	0.14	NB		1.1	0.15	5.1	0.08
0.4 to 0.6	2.5	0.03	2.2	0.03	10	0.41	NB		0.51	0.18	4.1	0.10
0.6 to 0.8	0.50	0.02	0.40	0.02	2.1	0.53	NB		0.077	0.20	0.96	0.07
0.8 to 1.0	0.24	0.02	0.16	0.03	0.61	0.45	NB		0.023	0.65	0.44	0.09
1.0 to 1.2	0.36	0.02	0.16	0.05	0.40	0.59	7.3	3.61	0.040	0.48	0.37	0.13
1.2 to 1.4	0.094	0.02	0.024	0.07	0.019	1.21	NP		0.0094	0.73	0.034	0.21

Table E.6 As Table E.1, LLN, binaries within 7.5", apparent radius of planets after blending within the pixel. For use with Fig. 8.4, lower panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	140	1.00	140	1.00	200	1.67	NB		64	1.51	190	1.67
-0.2 to 0	170	0.29	160	0.43	210	0.50	NB		110	0.50	180	0.40
0 to 0.2	52	0.13	51	0.13	100	0.15	NB		16	0.23	74	0.21
0.2 to 0.4	17	0.03	15	0.03	44	0.12	NB		4.0	0.13	23	0.08
0.4 to 0.6	4.1	0.03	3.8	0.04	20	0.41	NB		0.88	0.20	7.7	0.10
0.6 to 0.8	0.53	0.05	0.46	0.02	3.1	0.60	NB		0.10	0.26	1.3	0.10
0.8 to 1.0	0.25	0.03	0.20	0.04	1.1	0.64	NB		0.042	0.71	0.57	0.16
1.0 to 1.2	0.16	0.04	0.14	0.05	0.55	0.55	22	3.43	0.039	0.67	0.41	0.17
1.2 to 1.4	0.018	0.11	0.015	0.11	0.028	1.68	NB		0.0013	3.00	0.026	0.50
'omega'												
-0.4 to -0.2	6.6	1.03	6.2	1.08	12	1.71	NB		2.2	1.88	9.5	1.72
-0.2 to 0	13	0.29	12	0.43	15	0.52	NB		9.3	0.53	13	0.40
0 to 0.2	6.0	0.13	5.7	0.13	10	0.16	NB		1.8	0.30	8.3	0.21
0.2 to 0.4	3.9	0.03	3.7	0.04	9.7	0.14	NB		0.95	0.15	5.2	0.08
0.4 to 0.6	2.0	0.03	1.9	0.03	9.7	0.41	NB		0.41	0.20	3.8	0.10
0.6 to 0.8	0.40	0.01	0.35	0.01	2.1	0.53	NB		0.070	0.20	0.95	0.07
0.8 to 1.0	0.20	0.02	0.15	0.03	0.63	0.45	NB		0.024	0.40	0.45	0.09
1.0 to 1.2	0.14	0.03	0.12	0.04	0.39	0.63	7.3	3.61	0.039	0.34	0.34	0.14
1.2 to 1.4	0.022	0.05	0.018	0.06	0.019	1.21	NP		0.0053	1.00	0.031	0.14

Table E.7 As Table E.1, LLS, binaries within 7.5", true radius of planets. For use with Fig. 8.5, upper panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	240	1.60	230	1.60	290	2.00	NB		160	2.01	300	2.00
-0.2 to 0	190	0.15	180	0.15	430	0.64	NB		78	0.34	250	0.25
0 to 0.2	92	0.09	88	0.11	180	0.40	130	1.02	28	0.18	120	0.12
0.2 to 0.4	29	0.02	27	0.03	55	0.13	39	2.01	6.7	0.13	35	0.08
0.4 to 0.6	8.5	0.01	8.1	0.01	25	0.27	2000	2.24	2.0	0.21	13	0.13
0.6 to 0.8	1.1	0.05	0.91	0.05	3.0	0.59	12	3.00	0.22	0.31	1.7	0.12
0.8 to 1.0	0.51	0.02	0.33	0.02	0.95	0.43	2.6	2.83	0.075	0.51	0.61	0.14
1.0 to 1.2	0.56	0.02	0.28	0.04	0.49	0.39	NP		0.075	0.27	0.51	0.13
1.2 to 1.4	0.072	0.04	0.023	0.13	0.054	1.20	NP		0.0029	2.01	0.033	0.51
'omega'												
-0.4 to -0.2	15	1.63	15	1.63	17	2.03	NB		12	2.14	18	2.02
-0.2 to 0	13	0.16	13	0.16	34	0.65	NB		4.9	0.34	18	0.27
0 to 0.2	10	0.09	9.6	0.11	20	0.40	11	1.14	2.9	0.20	13	0.12
0.2 to 0.4	6.4	0.02	6.0	0.03	12	0.15	4.6	2.12	1.5	0.11	7.7	0.08
0.4 to 0.6	4.2	0.02	4.0	0.01	12	0.27	600	2.60	1.0	0.21	6.4	0.13
0.6 to 0.8	0.85	0.03	0.69	0.04	2.3	0.54	12	2.00	0.15	0.34	1.4	0.10
0.8 to 1.0	0.39	0.02	0.25	0.03	0.86	0.22	NP		0.055	0.30	0.47	0.08
1.0 to 1.2	0.47	0.01	0.23	0.02	0.49	0.52	0.28	2.24	0.042	0.29	0.44	0.08
1.2 to 1.4	0.087	0.04	0.028	0.03	0.063	0.63	NP		0.0043	1.34	0.052	0.23

Table E.8 As Table E.1, LLS, binaries within 7.5", apparent radius of planets after blending within the pixel. For use with Fig. 8.5, lower panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	240	1.60	230	1.60	290	2.00	NB		160	2.01	290	2.00
-0.2 to 0	200	0.15	190	0.15	450	0.64	NB		75	0.35	260	0.25
0 to 0.2	86	0.09	83	0.11	170	0.40	120	1.02	24	0.18	110	0.12
0.2 to 0.4	27	0.02	26	0.03	56	0.13	39	2.01	5.9	0.13	34	0.08
0.4 to 0.6	7.3	0.01	7.0	0.01	24	0.27	2000	2.24	1.7	0.20	12	0.13
0.6 to 0.8	0.96	0.05	0.82	0.05	3.0	0.59	12	3.00	0.20	0.34	1.7	0.12
0.8 to 1.0	0.43	0.04	0.33	0.04	0.97	0.43	2.6	2.83	0.065	0.67	0.64	0.12
1.0 to 1.2	0.28	0.02	0.21	0.04	0.48	0.39	NP		0.071	0.27	0.46	0.14
1.2 to 1.4	0.026	0.12	0.019	0.12	0.054	1.20	NP		0.0029	2.01	0.032	0.51
'omega'												
-0.4 to -0.2	15	1.63	15	1.63	17	2.03	NB		12	2.13	18	2.03
-0.2 to 0	14	0.16	13	0.15	35	0.64	NB		4.8	0.34	20	0.26
0 to 0.2	9.8	0.09	9.3	0.11	19	0.40	11	1.18	2.7	0.20	13	0.12
0.2 to 0.4	6.3	0.02	5.9	0.04	12	0.14	4.6	2.12	1.4	0.12	7.8	0.07
0.4 to 0.6	3.6	0.02	3.4	0.02	12	0.27	600	2.60	0.87	0.21	6.0	0.13
0.6 to 0.8	0.73	0.04	0.63	0.04	2.3	0.54	12	2.00	0.16	0.26	1.4	0.10
0.8 to 1.0	0.34	0.03	0.25	0.04	0.91	0.23	NP		0.051	0.47	0.50	0.07
1.0 to 1.2	0.24	0.02	0.19	0.02	0.46	0.60	0.28	2.24	0.039	0.29	0.41	0.11
1.2 to 1.4	0.030	0.04	0.023	0.03	0.063	0.63	NP		0.0043	1.34	0.050	0.23

Table E.9 As Table E.1, LLN, binaries within 45", true radius of planets. For use with Fig. 8.6, upper panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	4.2	0.16	4.2	0.16	8.1	0.20	9.5	1.03	2.1	0.30	6.6	0.16
-0.2 to 0	4.5	0.05	4.3	0.05	8.0	0.12	31	0.55	1.8	0.11	6.3	0.36
0 to 0.2	2.3	0.02	2.2	0.02	4.3	0.07	100	2.01	0.93	0.07	3.0	0.04
0.2 to 0.4	0.74	0.01	0.67	0.01	1.7	0.07	4.5	2.03	0.24	0.08	0.96	0.02
0.4 to 0.6	0.16	0.02	0.15	0.02	0.61	0.08	1.7	1.17	0.037	0.13	0.27	0.05
0.6 to 0.8	0.021	0.02	0.017	0.02	0.12	0.34	NB		0.0033	0.38	0.045	0.09
0.8 to 1.0	0.011	0.02	0.0074	0.03	0.095	0.50	NP		0.0012	0.67	0.028	0.15
1.0 to 1.2	0.017	0.02	0.0083	0.03	0.15	0.61	22	3.43	0.0025	0.50	0.039	0.13
1.2 to 1.4	0.0042	0.05	0.0013	0.09	0.011	1.67	NP		0.0001	3.00	0.0058	0.50
'omega'												
-0.4 to -0.2	0.20	0.28	0.19	0.38	0.48	0.41	0.71	1.17	0.067	0.97	0.33	0.43
-0.2 to 0	0.32	0.08	0.32	0.08	0.55	0.16	1.4	0.97	0.15	0.24	0.44	0.36
0 to 0.2	0.26	0.04	0.24	0.05	0.44	0.07	8.0	2.01	0.099	1.59	0.33	0.04
0.2 to 0.4	0.16	0.02	0.15	0.02	0.37	0.10	1.3	2.04	0.056	0.09	0.21	0.03
0.4 to 0.6	0.081	0.01	0.071	0.02	0.29	0.11	1.1	1.28	0.017	0.15	0.13	0.03
0.6 to 0.8	0.016	0.01	0.013	0.01	0.081	0.19	NB		0.0023	0.19	0.034	0.05
0.8 to 1.0	0.0085	0.02	0.0057	0.02	0.056	0.22	NP		0.0008	0.60	0.022	0.07
1.0 to 1.2	0.015	0.02	0.0071	0.05	0.11	0.55	22	3.61	0.0019	0.48	0.032	0.08
1.2 to 1.4	0.0051	0.01	0.0016	0.07	0.0071	1.21	NP		0.0007	0.73	0.0068	0.20

Table E.10 As Table E.1, LLN, binaries within 45", apparent radius of planets after blending within the pixel. For use with Fig. 8.6, lower panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	4.1	0.16	4.1	0.17	8.1	0.20	9.5	1.03	2.0	0.30	6.6	0.16
-0.2 to 0	4.6	0.05	4.4	0.05	8.3	0.12	29	0.58	1.7	0.11	6.5	0.36
0 to 0.2	2.2	0.02	2.0	0.02	4.1	0.07	98	2.01	0.79	0.07	2.8	0.03
0.2 to 0.4	0.68	0.01	0.62	0.01	1.7	0.07	4.5	2.03	0.21	0.05	0.95	0.02
0.4 to 0.6	0.13	0.01	0.12	0.02	0.58	0.10	1.7	1.17	0.030	0.017	0.25	0.05
0.6 to 0.8	0.017	0.05	0.015	0.02	0.12	0.34	NB		0.0031	0.25	0.044	0.09
0.8 to 1.0	0.0086	0.02	0.0070	0.03	0.099	0.50	NP		0.0015	0.67	0.028	0.15
1.0 to 1.2	0.0065	0.04	0.0060	0.04	0.15	0.51	22	3.43	0.0018	0.67	0.036	0.14
1.2 to 1.4	0.0010	0.11	0.0010	0.11	0.011	1.67	NP		0.0001	3.00	0.0053	0.50
'omega'												
-0.4 to -0.2	0.20	0.28	0.19	0.42	0.48	0.41	0.71	1.17	0.067	1.15	0.33	0.43
-0.2 to 0	0.34	0.08	0.34	0.08	0.59	0.18	1.4	0.97	0.15	0.21	0.47	0.36
0 to 0.2	0.25	0.03	0.23	0.04	0.42	0.08	8.0	2.01	0.091	0.21	0.32	0.04
0.2 to 0.4	0.16	0.02	0.14	0.03	0.38	0.10	1.4	2.05	0.049	0.09	0.21	0.04
0.4 to 0.6	0.066	0.01	0.060	0.02	0.28	0.09	0.92	1.22	0.017	0.17	0.12	0.04
0.6 to 0.8	0.013	0.01	0.011	0.01	0.080	0.18	NB		0.0021	0.19	0.033	0.05
0.8 to 1.0	0.0068	0.02	0.0055	0.02	0.057	0.21	NP		0.0008	0.32	0.022	0.07
1.0 to 1.2	0.0058	0.03	0.0053	0.04	0.11	0.60	7.3	3.61	0.0018	0.33	0.030	0.09
1.2 to 1.4	0.0012	0.05	0.0012	0.06	0.0071	1.21	NP		0.0004	1.00	0.0064	0.13

Table E.11 As Table E.1, LLS, binaries within 45", true radius of planets. For use with Fig. 8.7, upper panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	5.7	0.14	5.7	0.15	7.7	0.27	9.7	0.90	3.5	0.27	7.9	0.25
-0.2 to 0	7.8	0.06	7.6	0.06	12	0.16	30	0.93	3.6	0.09	11	0.12
0 to 0.2	4.3	0.02	4.1	0.03	7.9	0.09	66	1.02	1.7	0.09	5.3	0.04
0.2 to 0.4	1.4	0.01	1.3	0.01	3.3	0.10	12	1.02	0.47	0.07	1.9	0.02
0.4 to 0.6	0.31	0.01	0.28	0.02	1.0	0.13	1.6	1.05	0.076	0.10	0.50	0.04
0.6 to 0.8	0.041	0.04	0.033	0.04	0.17	0.32	0.51	3.61	0.0074	0.30	0.080	0.08
0.8 to 1.0	0.021	0.02	0.014	0.02	0.11	0.41	0.51	2.24	0.0034	0.50	0.043	0.12
1.0 to 1.2	0.030	0.01	0.016	0.03	0.13	0.33	NP		0.0046	0.25	0.062	0.13
1.2 to 1.4	0.0060	0.04	0.0023	0.13	0.026	1.22	NP		0.0004	2.00	0.0082	0.50
'omega'												
-0.4 to -0.2	0.36	0.30	0.36	0.31	0.44	0.42	0.33	0.80	0.26	0.78	0.47	0.38
-0.2 to 0	0.55	0.09	0.54	0.08	0.96	0.19	1.9	0.92	0.23	0.08	0.76	0.14
0 to 0.2	0.48	0.03	0.44	0.03	0.87	0.09	5.8	1.14	0.18	0.13	0.58	0.04
0.2 to 0.4	0.32	0.02	0.29	0.02	0.71	0.12	1.4	1.22	0.11	0.04	0.41	0.02
0.4 to 0.6	0.15	0.02	0.14	0.02	0.49	0.13	0.47	1.69	0.039	0.11	0.24	0.05
0.6 to 0.8	0.031	0.02	0.025	0.02	0.13	0.22	0.51	2.83	0.0052	0.33	0.0066	0.03
0.8 to 1.0	0.016	0.02	0.011	0.03	0.10	0.17	NP		0.0025	0.29	0.033	0.06
1.0 to 1.2	0.025	0.01	0.014	0.02	0.13	0.48	0.28	2.24	0.0026	0.27	0.054	0.08
1.2 to 1.4	0.0073	0.04	0.0028	0.03	0.030	0.67	NP		0.0005	1.33	0.013	0.22

Table E.12 As Table E.1, LLS, binaries within 45", apparent radius of planets after blending within the pixel. For use with Fig. 8.7, lower panel.

Radius/ *log* $R_\oplus$	$V < 26$	$\delta X/X$	$V < 16$, F5-M	$\delta X/X$	P1	$\delta X/X$	P2	$\delta X/X$	P4	$\delta X/X$	P5	$\delta X/X$
'zeta'												
-0.4 to -0.2	5.7	0.14	5.5	0.14	7.7	0.27	9.7	0.90	3.4	0.32	7.8	0.25
-0.2 to 0	8.0	0.06	7.9	0.06	13	0.16	30	0.93	3.5	0.12	11	0.12
0 to 0.2	4.0	0.02	3.8	0.03	7.6	0.09	64	1.02	1.5	0.08	5.1	0.04
0.2 to 0.4	1.3	0.01	1.2	0.02	3.3	0.10	12	1.02	0.42	0.08	1.8	0.03
0.4 to 0.6	0.27	0.01	0.24	0.01	1.0	0.14	1.6	1.05	0.066	0.08	0.47	0.04
0.6 to 0.8	0.035	0.04	0.030	0.04	0.17	0.32	0.51	3.61	0.0070	0.33	0.080	0.08
0.8 to 1.0	0.018	0.03	0.014	0.03	0.12	0.41	0.51	2.24	0.0030	0.67	0.045	0.10
1.0 to 1.2	0.015	0.02	0.013	0.04	0.13	0.33	NP		0.0046	0.25	0.057	0.15
1.2 to 1.4	0.0022	0.12	0.0019	0.12	0.026	1.22	NP		0.0004	2.00	0.0080	0.50
'omega'												
-0.4 to -0.2	0.36	0.30	0.36	0.31	0.44	0.42	0.33	1.27	0.25	0.77	0.48	0.39
-0.2 to 0	0.59	0.09	0.57	0.07	1.0	0.19	2.0	0.93	0.23	0.11	0.82	0.14
0 to 0.2	0.46	0.04	0.43	0.04	0.85	0.10	5.5	1.18	0.17	0.11	0.56	0.05
0.2 to 0.4	0.31	0.02	0.28	0.02	0.72	0.11	1.4	1.22	0.10	0.05	0.42	0.02
0.4 to 0.6	0.13	0.02	0.12	0.02	0.47	0.13	0.47	1.69	0.034	0.11	0.23	0.05
0.6 to 0.8	0.027	0.02	0.023	0.02	0.13	0.22	0.51	2.83	0.0054	0.25	0.064	0.03
0.8 to 1.0	0.014	0.03	0.011	0.03	0.11	0.18	NP		0.0023	0.46	0.035	0.03
1.0 to 1.2	0.013	0.02	0.011	0.02	0.12	0.56	0.28	2.24	0.0024	0.28	0.050	0.11
1.2 to 1.4	0.0025	0.04	0.0022	0.03	0.030	0.67	NP		0.0005	1.33	0.012	0.22

Table E.13 Ratio of main sequence single stars and binaries where the evolutionary primary is on the main sequence, integrated along the line of sight, LLN to LLS, for comparison with Table 8.4. $|b| = 6^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, $V \leq 26$.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	1/4 = 0.25	1/1 = 1.00	0/1 = 0	4/1 = 4.00	1/1 = 1.00
0.10	6/12 = 0.5	6/5 = 1.20	3/6 = 0.5	14/7 = 2.00	6/7 = 0.86
0.20	59/67 = 0.88	52/55 = 0.95	62/65 = 0.95	96/80 = 1.20	65/75 = 0.87
0.30	171/218 = 0.78	207/188 = 1.10	196/228 = 0.86	293/256 = 1.14	229/226 = 1.01
0.40	445/500 = 0.89	513/440 = 1.17	499/553 = 0.90	678/618 = 1.10	559/581 = 0.96
0.50	893/957 = 0.93	978/863 = 1.13	951/1095 = 0.87	1353/1160 = 1.17	1168/1125 = 1.04
0.60	1507/1609 = 0.94	1654/1477 = 1.12	1639/1856 = 0.88	2290/2209 = 1.14	2031/1930 = 1.05
0.70	2335/2452 = 0.95	2603/2276 = 1.14	2531/2803 = 0.90	3640/3160 = 1.15	3262/2929 = 1.11
0.80	3421/3552 = 0.96	3800/3284 = 1.16	3716/4052 = 0.92	5349/4598 = 1.16	4866/4319 = 1.13
0.90	4770/4863 = 0.98	5283/4493 = 1.18	5166/5562 = 0.93	7545/6372 = 1.18	6936/6044 = 1.15
1.00	6406/6405 = 1.00	7071/5946 = 1.19	6945/7331 = 0.95	10160/8453 = 1.20	9437/8134 = 1.16
2.00	35165/30424 = 1.16	39266/31258 = 1.26	42453/37258 = 1.14	55417/41509 = 1.34	58310/45390 = 1.29
3.00	78650/61452 = 1.28	89795/66532 = 1.35	104862/77595 = 1.35	114388/82584 = 1.39	141910/101808 = 1.39
4.00	124147/89458 = 1.39	146659/101780 = 1.44	183122/115669 = 1.58	170858/118713 = 1.44	247611/163823 = 1.51
5.00	162196/110429 = 1.47	201605/132621 = 1.52	266015/144442 = 1.84	220667/142794 = 1.55	363198/221451 = 1.64
6.00	191884/125174 = 1.53	249150/157076 = 1.59	345142/163232 = 2.11	264861/159079 = 1.67	474528/270206 = 1.76
7.00	214517/135225 = 1.59	286494/175824 = 1.63	415345/176255 = 2.36	303286/171227 = 1.77	574516/307821 = 1.87
8.00	230314/142092 = 1.62	314195/189896 = 1.66	471987/185544 = 2.54	335265/180267 = 1.86	657807/335846 = 1.96
9.00	241832/146687 = 1.65	334514/199742 = 1.68	514217/192035 = 2.68	360605/186786 = 1.93	724890/355798 = 2.04
10.00	250185/149785 = 1.67	349246/206607 = 1.69	545852/196723 = 2.77	380217/191623 = 1.98	776359/369963 = 2.10

Table E.14 As Table E.13, $|b| = 30^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, $V \leq 26$.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	1/1 = 1.00	0/0	1/2 = 0.50	1/3 = 0.33	1/0
0.10	3/6 = 0.50	5/3 = 1.67	5/6 = 0.83	8/6 = 1.33	3/7 = 0.43
0.20	32/46 = 0.70	34/39 = 0.87	44/51 = 0.86	47/50 = 0.94	39/33 = 1.18
0.30	115/135 = 0.85	123/138 = 0.89	124/145 = 0.86	111/139 = 0.80	116/127 = 0.91
0.40	233/296 = 0.79	258/280 = 0.92	262/286 = 0.92	242/299 = 0.81	241/283 = 0.85
0.50	431/506 = 0.85	456/461 = 0.99	459/485 = 0.95	435/525 = 0.83	442/496 = 0.89
0.60	701/776 = 0.90	714/720 = 0.99	733/793 = 0.92	711/793 = 0.90	716/749 = 0.96
0.70	990/1068 = 0.93	1022/028 = 0.99	1060/1086 = 0.98	1049/1115 = 0.94	1042/1073 = 0.97
0.80	1307/1390 = 0.94	1350/1362 = 0.99	1430/1482 = 0.97	1387/1483 = 0.94	1467/1489 = 0.99
0.90	1662/1758 = 0.95	1763/1720 = 1.03	1852/1898 = 0.98	1806/1890 = 0.96	1916/1939 = 0.99
1.00	2031/2129 = 0.95	2187/2119 = 1.03	2334/2354 = 0.99	2276/2349 = 0.97	2422/2433 = 1.00
2.00	6521/6127 = 1.06	7099/6369 = 1.12	7662/6826 = 1.12	8027/7377 = 1.09	8639/7881 = 1.10
3.00	10505/9032 = 1.16	11643/9665 = 1.21	12832/10455 = 1.23	13890/11465 = 1.21	15177/12480 = 1.22
4.00	13724/11138 = 1.23	15482/12053 = 1.28	17239/13269 = 1.30	19266/14708 = 1.31	21583/16264 = 1.33
5.00	16182/12655 = 1.28	18551/13840 = 1.34	21154/15278 = 1.39	24173/17016 = 1.42	27362/19140 = 1.43
6.00	18060/13639 = 1.32	20999/15053 = 1.40	24405/16695 = 1.46	28221/18709 = 1.51	32649/21458 = 1.52
7.00	19331/14307 = 1.35	22729/15862 = 1.43	26721/17736 = 1.51	31368/19859 = 1.58	36899/23015 = 1.60
8.00	20211/14735 = 1.37	23875/16340 = 1.46	28279/18334 = 1.54	33556/20606 = 1.63	39972/24051 = 1.66
9.00	20757/14984 = 1.39	24610/16637 = 1.48	29307/18687 = 1.57	35055/21083 = 1.66	42289/24709 = 1.71
10.00	21085/15140 = 1.39	25066/16781 = 1.49	29991/18900 = 1.59	36104/21361 = 1.69	43867/25075 = 1.75

Table E.15 As Table E.13, $|b| = 54^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, $V \leq 26$.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/1 = 0.00	0/1 = 0.00	0/0	0/04
0.10	5/3 = 1.67	4/1 = 4.00	4/4 = 1.00	1/4 = 0.25	4/6 = 0.67
0.20	24/22 = 1.09	41/18 = 2.28	26/23 = 1.13	22/22 = 1.00	30/28 = 1.07
0.30	57/67 = 0.85	73/79 = 0.92	65/72 = 0.90	66/60 = 1.10	86/74 = 1.16
0.40	125/135 = 0.93	146/149 = 0.98	118/131 = 0.90	118/142 = 0.83	153/150 = 1.02
0.50	225/242 = 0.93	235/237 = 0.99	196/230 = 0.85	206/240 = 0.86	226/266 = 0.85
0.60	328/338 = 0.97	340/348 = 0.98	320/359 = 0.89	303/368 = 0.82	235/403 = 0.83
0.70	429/453 = 0.95	451/464 = 0.97	440/490 = 0.90	401/486 = 0.83	457/534 = 0.86
0.80	565/590 = 0.96	563/624 = 0.90	560/649 = 0.86	517/635 = 0.81	598/668 = 0.90
0.90	694/732 = 0.95	689/752 = 0.92	683/767 = 0.89	670/792 = 0.85	740/821 = 0.90
1.00	816/878 = 0.93	821/899 = 0.91	844/900 = 0.94	804/961 = 0.84	892/989 = 0.90
2.00	2024/2000 = 1.01	2099/2103 = 1.00	2170/2150 = 1.01	2260/2250 = 1.00	2331/2339 = 1.00
3.00	2963/2859 = 1.04	3162/2889 = 1.09	3273/3058 = 1.07	3506/3226 = 1.09	3658/3407 = 1.07
4.00	3729/3365 = 1.11	3976/3484 = 1.14	4212/3762 = 1.12	4482/3973 = 1.13	4754/4235 = 1.12
5.00	4190/3677 = 1.14	4491/3859 = 1.16	4867/4162 = 1.17	5220/4475 = 1.17	5574/4751 = 1.17
6.00	4471/3871 = 1.16	4806/4063 = 1.18	5260/4417 = 1.19	5676/4750 = 1.20	6131/5079 = 1.21
7.00	4623/2980 = 1.16	4995/4172 = 1.20	5502/4523 = 1.22	5953/4887 = 1.22	6475/5255 = 1.23
8.00	4695/4024 = 1.17	5086/4217 = 1.21	5616/4593 = 1.22	6125/4956 = 1.24	6668/5343 = 1.25
9.00	4727/4039 = 1.17	5132/4235 = 1.21	5691/4614 = 1.23	6212/4984 = 1.25	6769/5385 = 1.26
10.00	4748/4046 = 1.17	5147/4243 = 1.21	5723/4628 = 1.24	6254/5005 = 1.25	6825/5408 = 1.26

Table E.16 As Table E.13, $|b| = 6^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, $V \leq 16$, F5-M. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/2 = 0.00	0/0	0/0	2/1 = 2.00	0/1 = 0.00
0.10	2/7 = 0.29	2/2 = 1.00	3/2 = 1.5	7/3 = 2.33	3/3 = 1.00
0.20	13/25 = 0.52	11/12 = 0.92	20/14 = 1.43	19/16 = 1.19	14/11 = 1.27
0.30	23/46 = 0.50	33/34 = 0.97	37/33 = 1.12	45/29 = 1.55	31/24 = 1.29
0.40	51/81 = 0.63	55/65 = 0.87	55/71 = 0.78	65/57 = 1.14	59/48 = 1.23
0.50	84/122 = 0.69	89/111 = 0.80	95/109 = 0.87	113/102 = 1.11	100/90 = 1.11
0.60	125/168 = 0.74	129/156 = 0.83	133/171 = 0.78	154/144 = 1.07	147/142 = 1.04
0.70	168/228 = 0.74	175/223 = 0.78	187/225 = 0.83	219/197 = 1.11	203/194 = 1.05
0.80	222/284 = 0.78	242/282 = 0.86	243/301 = 0.81	269/257 = 1.05	275/277 = 0.99
0.90	275/360 = 0.76	302/348 = 0.87	316/360 = 0.88	321/303 = 1.06	333/360 = 0.93
1.00	329/429 = 0.77	376/435 = 0.86	404/444 = 0.91	377/361 = 1.04	420/452 = 0.93
2.00	826/893 = 0.92	833/1117 = 0.75	1049/945 = 1.11	673/668 = 1.01	972/1062 = 0.92
3.00	944/967 = 0.98	928/1381 = 0.67	1264/1037 = 1.22	682/690 = 0.99	1038/1201 = 0.86
4.00	057/974 = 0.98	935/1428 = 0.65	1314/1049 = 1.25	682/690 = 0.99	1047/1223 = 0.86
5.00	957/974 = 0.98	937/1434 = 0.65	1331/1049 = 1.27	682/690 = 0.99	1047/1223 = 0.86
6.00	957/974 = 0.98	937/1435 = 0.65	1331/1049 = 1.27	682/690 = 0.99	1047/1223 = 0.86

Table E.17 As Table E.13, $|b| = 30^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, $V \leq 16$, F5-M. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	1/1 = 1.00	0/0	0/2	1/3 = 0.33	1/0
0.10	2/5 = 0.40	2/2 = 1.00	1/3 = 0.33	3/3 = 1.00	1/4 = 0.25
0.20	9/14 = 0.64	12/7 = 1.71	12/19 = 0.63	10/14 = 0.71	13/15 = 0.87
0.30	27/29 = 0.93	25/25 = 1.00	19/34 = 0.56	23/23 = 1.00	25/32 = 0.78
0.40	35/53 = 0.66	39/46 = 0.85	35/52 = 0.67	40/40 = 1.00	34/49 = 0.69
0.50	53/77 = 0.69	61/68 = 0.90	56/69 = 0.81	63/57 = 1.11	54/69 = 0.78
0.60	81/96 = 0.84	85/96 = 0.89	82/92 = 0.89	84/80 = 1.05	69/88 = 0.78
0.70	105/126 = 0.83	110/116 = 0.95	106/115 = 0.92	106/108 = 0.98	92/112 = 0.82
0.80	130/146 = 0.89	134/142 = 0.94	142/150 = 0.95	123/129 = 0.95	125/135 = 0.93
0.90	140/173 = 0.81	156/168 = 0.93	170/172 = 0.99	147/158 = 0.93	151/166 = 0.91
1.00	169/193 = 0.88	182/192 = 0.95	198/195 = 1.02	172/184 = 0.93	183/191 = 0.96
2.00	302/326 = 0.93	322/346 = 0.93	335/344 = 0.97	364/348 = 1.05	388/350 = 1.11
3.00	327/356 = 0.92	357/367 = 0.97	372/372 = 1.00	395/373 = 1.06	420/376 = 1.12
4.00	329/357 = 0.93	361/370 = 0.98	374/373 = 1.00	396/373 = 1.06	423/377 = 1.12

Table E.18 As Table E.13, $|b| = 54^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, $V \leq 16$, F5-M. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/0	0/0	0/0
0.10	1/1 = 1.00	2/0	3/1 = 3.00	0/2 = 0.00	2/1 = 2.00
0.20	5/7 = 0.71	10/7 = 1.43	11/4 = 2.75	8/6 = 1.33	8/8 = 1.00
0.30	11/14 = 0.79	15/16 = 0.94	17/14 = 1.21	12/15 = 0.80	20/18 = 1.11
0.40	17/24 = 0.71	27/25 = 1.08	22/27 = 0.81	20/21 = 0.95	29/30 = 0.97
0.50	28/35 = 0.80	36/33 = 1.09	31/35 = 0.89	31/28 = 1.11	33/42 = 0.79
0.60	34/48 = 0.71	50/42 = 1.19	51/47 = 1.09	39/42 = 0.93	42/58 = 0.72
0.70	45/58 = 0.78	61/55 = 1.11	61/51 = 1.20	48/54 = 0.89	51/73 = 0.70
0.80	54/69 = 0.78	76/62 = 1.23	65/59 = 1.10	59/66 = 0.89	63/85 = 0.74
0.90	67/79 = 0.85	82/71 = 1.16	71/68 = 1.04	67/80 = 0.84	73/97 = 0.75
1.00	75/85 = 0.88	94/76 = 1.24	80/80 = 1.00	71/90 = 0.79	85/108 = 0.79
2.00	116/119 = 0.98	136/104 = 1.31	122/117 = 1.04	98/126 = 0.78	124/148 = 0.84
3.00	120/123 = 0.98	143/109 = 1.31	128/123 = 104	100/129 = 0.78	129/152 = 0.85
4.00	120/123 = 0.98	143/109 = 1.31	128/123 = 104	101/129 = 0.78	129/152 = 0.85

Table E.19 As Table E.13, $|b| = 6^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, P1. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/0	0/0	0/0
0.10	1/1 = 1.00	0/0	1/0	1/0	0/0
0.20	1/3 = 0.33	1/2 = 0.50	1/2 = 0.50	3/1 = 3.00	1/3 = 0.33
0.30	1/4 = 0.25	1/3 = 0.33	1/3 = 0.33	4/1 = 4.00	3/4 = 0.75
0.40	3/7 = 0.43	2/4 = 0.50	2/4 = 0.50	5/2 = 2.50	4/4 = 1.00
0.50	3/8 = 0.38	2/4 = 0.50	2/4 = 0.50	5/2 = 2.50	4/4 = 1.00
0.60	3/8 = 0.38	2/4 = 0.50	2/4 = 0.50	5/2 = 2.50	4/4 = 1.00
0.70	3/8 = 0.38	2/5 = 0.40	2/4 = 0.50	6/2 = 3.00	4/4 = 1.00

Table E.20 As Table E.13, $|b| = 30^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, P1. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/1 = 0.00	0/2 = 0.00	0/0
0.10	0/0	0/0	0/1 = 0.00	0/2 = 0.00	0/1 = 0.00
0.20	2/0	2/1 = 2.00	0/2 = 0.00	0/3 = 0.00	0/2 = 0.00
0.30	3/0	2/1 = 2.00	0/2 = 0.00	0/3 = 0.00	0/3 = 0.00
0.40	3/0	2/2 = 1.00	0/2 = 0.00	0/3 = 0.00	1/3 = 0.33
0.50	3/0	2/2 = 1.00	0/2 = 0.00	0/3 = 0.00	1/3 = 0.33
0.60	3/0	2/2 = 1.00	0/3 = 0.00	0/3 = 0.00	1/3 = 0.33

Table E.21 As Table E.13, $|b| = 54^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, P1. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/0	0/0	0/0
0.10	0/0	0/0	0/0	0/0	0/0
0.20	0/1 = 0.00	1/1 = 1.00	2/0	2/0	0/0
0.30	1/1 = 1.00	1/2 = 0.50	2/1 = 2.00	3/1 = 3.00	0/1 = 0.00
0.40	1/2 = 0.50	2/2 = 1.00	2/1 = 2.00	3/1 = 3.00	0/1 = 0.00
0.50	1/2 = 0.50	2/3 = 0.67	2/1 = 2.00	3/2 = 1.50	0/1 = 0.00
0.60	1/2 = 0.50	2/3 = 0.67	2/1 = 2.00	3/2 = 1.50	0/1 = 0.00
0.70	1/2 = 0.50	2/3 = 0.67	2/1 = 2.00	3/2 = 1.50	1/1 = 1.00

Table E.22 As Table E.13, $|b| = 6° \pm 0.5°$, $l \pm 0.5°$, P2. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/0	0/0	0/0

Table E.23 As Table E.13, $|b| = 30° \pm 0.5°$, $l \pm 0.5°$, P2. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/0	0/1 = 0.00	0/0

Table E.24 As Table E.13, $|b| = 54° \pm 0.5°$, $l \pm 0.5°$, P2. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/0	0/0	0/0

Table E.25 As Table E.13, $|b| = 6° \pm 0.5°$, $l \pm 0.5°$, P4. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/2	0/0	0/1 = 0	2/1 = 2.00	0/1 = 0.00
0.10	1/4 = 0.25	1/1 = 1.00	1/2 = 0.5	4/2 = 2.00	3/2 = 1.50
0.20	6/12 = 0.50	2/2 = 1.00	7/2 = 0.5	7/3 = 2.33	7/2 = 3.50

Table E.26 As Table E.13, $|b| = 30° \pm 0.5°$, $l \pm 0.5°$, P4. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	1/1 = 1.00	0/0	0/1 = 0.00	1/1 = 1.00	1/0
0.10	1/5 = 0.20	1/1 = 1.00	1/2 = 0.50	2/1 = 2.00	1/1 = 1.00
0.20	4/9 = 0.44	6/1 = 6.00	8/7 = 1.14	3/3 = 1.00	5/4 = 1.25

Table E.27 As Table E.13, $|b| = 54^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, P4. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/0	0/0	0/0
0.10	1/1 = 1.00	2/0	2/1 = 2.00	0/1 = 0.00	1/1 = 1.00
0.20	3/1 = 3.00	5/2 = 2.5	5/2 = 2.50	3/3 = 1.00	3/5 = 0.60
0.30	3/1 = 3.00	5/2 = 2.5	5/2 = 2.50	3/3 = 1.00	4/5 = 0.67

Table E.28 As Table E.13, $|b| = 6^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, P5. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/0	0/0	0/0
0.10	1/3 = 0.33	0/0	2/0	3/1 = 3	0/1 = 0.00
0.20	2/10 = 0.20	5/5 = 1.00	9/7 = 1.29	7/6 = 1.2	4/6 = 0.67
0.30	5/17 = 0.29	11/11 = 1.00	12/11 = 1.09	15/8 = 1.88	8/14 = 0.57
0.40	15/25 = 0.60	20/20 = 1.00	16/23 = 0.70	20/16 = 1.25	17/20 = 0.85
0.50	20/39 = 0.51	26/31 = 0.84	20/29 = 0.69	30/30 = 1.00	22/26 = 0.85
0.60	25/46 = 0.54	33/42 = 0.79	22/33 = 0.67	34/37 = 0.92	27/29 = 0.93
0.70	27/48 = 0.56	36/48 = 0.75	31/37 = 0.84	38/41 = 0.93	27/32 = 0.84
0.80	28/50 = 0.56	38/48 = 0.75	31/38 = 0.82	39/43 = 0.91	29/36 = 0.81
0.90	30/50 = 0.60	38/53 = 0.72	32/39 = 0.82	40/45 = 0.89	29/39 = 0.74
1.00	30/53 = 0.57	40/55 = 0.73	34/41 = 0.83	41/45 = 0.91	30/39 = 0.77
2.00	30/56 = 0.54	42/57 = 0.74	36/43 = 0.84	41/46 = 0.89	31/41 = 0.76

Table E.29 As Table E.13, $|b| = 30^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, P5. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/1 = 0.00	0/2 = 0.00	0/0
0.10	0/0	1/1 = 1.00	0/1 = 0.00	1/2 = 0.50	0/1 = 0.00
0.20	2/3 = 0.67	5/5 = 1.00	1/8 = 0.13	3/5 = 0.60	6/6 = 1.00
0.30	9/7 = 1.29	9/11 = 0.82	4/17 = 0.24	6/8 = 0.75	10/10 = 1.00
0.40	12/12 = 1.00	13/13 = 1.00	7/24 = 0.29	12/10 = 1.2	12/16 = 0.75
0.50	18/17 = 1.06	17/17 = 1.00	9/26 = 0.35	15/14 = 1.07	16/22 = 0.73
0.60	25/19 = 1.32	17/22 = 0.77	13/32 = 0.41	18/17 = 1.06	21/23 = 0.91
0.70	28/21 = 1.33	20/25 = 0.80	16/37 = 0.43	21/20 = 1.05	21/24 = 0.88
0.80	30/22 = 1.37	20/27 = 0.74	18/40 = 0.45	22/22 = 1.00	22/24 = 0.92
0.90	30/22 = 1.37	20/27 = 0.74	20/43 = 0.46	24/22 = 1.09	22/24 = 0.92
1.00	30/25 = 1.20	21/27 = 0.78	21/44 = 0.48	24/23 = 1.04	22/24 = 0.92
2.00	31/29 = 1.07	22/28 = 0.79	21/46 = 0.46	25/24 = 1.04	22/24 = 0.92

Table E.30 As Table E.13, $|b| = 54^\circ \pm 0.5^\circ$, $l \pm 0.5^\circ$, P5. No change in stellar numbers beyond last line in the table.

d/kpc	89° (LLN)/ 229° (LLS)	77° (LLN)/ 241° (LLS)	65° (LLN)/ 253° (LLS)	53° (LLN)/ 265° (LLS)	41° (LLN)/ 277° (LLS)
0.05	0/0	0/0	0/0	0/0	0/0
0.10	0/0	0/0	1/0	0/1 = 0.00	1/0
0.20	1/3 = 0.33	2/4 = 0.50	6/1 = 6.00	2/2 = 1.00	2/1 = 2.00
0.30	4/3 = 1.33	4/8 = 0.50	7/3 = 2.33	3/6 = 0.50	3/4 = 0.75
0.40	5/4 = 1.25	7/12 = 0.58	8/8 = 1.00	6/7 = 0.86	8/5 = 1.60
0.50	7/5 = 1.40	9/16 = 0.56	12/10 = 1.2	8/8 = 1.00	8/7 = 1.14
0.60	7/9 = 0.78	11/17 = 0.65	16/13 = 1.23	9/10 = 0.90	9/8 = 1.13
0.70	9/11 = 0.82	11/19 = 0.58	18/13 = 1.39	10/10 = 1.00	11/9 = 1.22
0.80	9/12 = 0.75	11/19 = 0.58	19/4 = 1.36	11/11 = 1.00	12/10 = 1.20
0.90	9/12 = 0.75	11/20 = 0.55	19/4 = 1.36	11/13 = 0.85	12/11 = 1.09
1.00	9/12 = 0.75	11/21 = 0.52	19/4 = 1.36	11/13 = 0.85	12/11 = 1.09
2.00	9/13 = 0.69	11/21 = 0.52	19/4 = 1.36	11/13 = 0.85	12/11 = 1.09